PHYSICAL REVIEW D VOLUME 19, NUMBER 2 15 JANUARY 1979

Quantum electrodynamics on a lattice: A Hamiltonian variational approach to the physics of the weak-coupling region

Sidney D. Drell, Helen R. Quinn, Benjamin Svetitsky, and Marvin Weinstein
Stanford Linear Accelerator Center, Stanford University, Stanford, California 94305
(Received 9 June 1978)

We develop and apply a Hamiltonian variational approach to the study of quantum electrodynamics formulated on a spatial lattice in both $2+1$ and $3+1$ dimensions. Two lattice versions of QED are considered: a noncompact version which reproduces the physics of continuum QED, and a compact version constructed in correspondence with lattice formulations of non-Abelian theories. Our focus is on photon dynamics with charged particles treated in the static limit. We are especially interested in the nonperturbative aspects of the solutions in the weak-coupling region in order to clarify and establish aspects of confinement. In particular we find, in accord with Polyakov, that the compact QED leads to linear confinement for any nonvanishing coupling, no matter how small, in $2+1$ dimensions, but that a phase transition to an unconfined phase for sufficiently weak couplings occurs in $3+1$ dimensions. We identify and describe the causes of confinement.

I. INTRODUCTION

This paper presents a study of two different lattice versions of quantum electrodynamics (QED) in both $2+1$ and $3+1$ dimensions. The focus of this work is on photon dynamics with the charged particles treated in the static limit. In this limit, the question of whether or not the theory exhibits linear confinement is reduced to computing the ground-state energy of the electrodynamic field theory in the presence of a pair of opposite charges separated by a distance D. Variational methods developed earlier[1] are used to estimate this ground-state energy.

Our study of QED is a step towards the goal of using nonperturbative variational methods to analyze confinement and other fundamental properties of non-Abelian gauge theories. We start with an analysis of the simpler Abelian theories in order to learn how to handle the additional constraints on the states imposed by gauge invariance, and in order to ascertain the extent to which the requirement of gauge invariance restricts the dynamical structure of the theory. As we will see, Abelian QED admits many inequivalent Hamiltonian formulations. In order to illustrate how changes in the Hamiltonian that are apparently minor can lead to major changes in physics, we analyze two specific models. The first model is constructed to reproduce the physics of continuum QED, which has noninteracting photons; this is called the noncompact version. The second version of QED is defined in correspondence with lattice formulations that have been heretofore constructed for non-Abelian theories,[2] such as the quantum chromodynamics (QCD) of quarks and gluons. This is the sort of theory which would arise naturally for QED if one started with a unified gauge theory of weak and electromagnetic interactions and identified the photon with the gauge field of an unbroken one-dimensional subgroup of the larger theory.[3] The version of QED which emerges in this way inherits from the larger theory the fact that photons are self-interacting through a potential which is a bounded periodic function of the photon field. It is this version of the theory which has been presented in the work of Wilson and of Kogut and Susskind[2] and which is referred to as compact, for reasons which will soon become clear. It is now well known that, for sufficiently strong coupling, compact lattice theories, Abelian or non-Abelian, exhibit linear confinement in both $2+1$ and $3+1$ dimensions.[2,3] Hence, whether or not they can provide a satisfactory formulation of QED depends on their behavior for small couplings.

In $2+1$ dimensions we find that the linear confinement persists for all nonvanishing couplings $g^2>0$, no matter how weak. This is in agreement with results obtained by Polyakov[3] who demonstrated it in $2+1$ dimensions for continuum SO(3) theory by path-integral methods and argued that it should also occur in a lattice self-interacting Abelian theory.

In $3+1$ dimensions we find a different situation in that for weak coupling, $g^2 \ll 1$, there is no linear confinement and the interaction between charges is Coulomb type. This result has also been obtained by Banks, Myerson, and Kogut[4] who used a Villain approximation to the Abelian lattice theory and argued by analogy to the work of Polyakov.

Our analysis shows that confinement in the compact theory is directly attributable to the fact that in this form of the theory photon self-interactions have been introduced. The reason these self-inter-

actions can lead to confinement can be qualitatively understood if we consider creation of a particle-antiparticle pair. When a pair of charges is created at two separated lattice points, causality demands that initially the electric field is confined to some small region in the neighborhood of the charges. This field can be interpreted as the Coulomb field of the separated charges plus a coherent cloud of transverse photons which cancel the Coulomb field everywhere except near the charges. In the noncompact version of the theory, as in the continuum case, the photons do not interact with each other, and they simply radiate away to infinity leaving the ground state of the system—the Coulomb field. In the compact formulation, for strong coupling, the photons interact so strongly among themselves that the coherent photon state is an approximate eigenstate of the system and does not radiate away. Hence the two charges remain joined by a tube of flux. At weaker couplings quantum fluctuations occur, and these randomize the coherent state to some extent. As we will demonstrate, in two spatial dimensions this randomization can never be sufficient to completely destroy this tube and one obtains linear confinement. In three spatial dimensions we show that for weak couplings the fluctuations become large enough to wipe out the tube of flux, and confinement does not survive. This greater randomization can be understood as a consequence of the fact that there is one more dimension in which the quantum fluctuations can occur. (It is interesting to note in connection with this picture of confinement, that non-Abelian gauge theories automatically describe self-interacting gauge fields.)

The next section of this paper gives our notation for and formulation of lattice QED in the $A_0=0$ gauge. In Sec. III we discuss in some detail the physics of $(2+1)$-dimensional QED in a very small universe—namely a single square of the lattice. We will be able to present most of our techniques in the context of this very simple problem, and also to demonstrate much of the physics discussed above. Section IV will then be devoted to the generalization of this treatment to a larger lattice, thus completing our discussion of $2+1$ dimensions. In Sec. V we present the analysis of $3+1$ dimensions, and in Sec. VI we summarize and speculate.

II. FORMULATION OF QED ON A LATTICE

A. Latticization

We present our analysis in the $A_0=0$ gauge because it is best suited to our Hamiltonian approach. For simplicity we describe our notation in terms of a two-dimensional planar lattice. Its extension to a cubic lattice in three space dimensions is straightforward. We define the canonically conjugate vector fields $\vec{A}$ and $\vec{E}$ by placing them on the links of the lattice, so that each site $\vec{p}=(p_x,p_y)$ has associated with it field components $A^x_{\vec{p}}$ and $A^y_{\vec{p}}$ residing on the links leaving the site in $+x$ and $+y$ directions, respectively (see Fig. 1). We then define $B_{\vec{p}}$ as the lattice curl

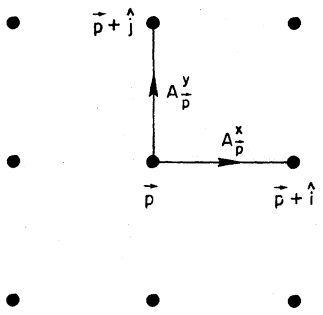

FIG. 1. Labeling of sites and links in a two-dimensional lattice.

$$B_{\vec{p}}=(\vec{\nabla}\times\vec{A})_{\vec{p}}\equiv\frac{1}{a}\left(A^x_{\vec{p}}+A^y_{\vec{p}+\hat{i}}-A^x_{\vec{p}+\hat{j}}-A^y_{\vec{p}}\right), \tag{2.1}$$

where a is the lattice spacing. Like all pseudovectors B has one component, directed out of the plane in accord with a right-hand rule and located at plaquette centers; plaquettes are labeled by their lower left-hand corners (Fig. 2). We also define the divergence of a vector (a scalar defined at each site; see Fig. 3)

$$(\vec{\nabla}\cdot\vec{E})_{\vec{p}}=\frac{1}{a}\left(E^x_{\vec{p}}+E^y_{\vec{p}}-E^x_{\vec{p}-\hat{i}}-E^y_{\vec{p}-\hat{j}}\right). \tag{2.2}$$

The Hamiltonian for the photon field in the $A_0=0$ gauge can be latticized as

$$H^{(1)}=\tfrac{1}{2}\int d^2x(E^2+B^2)$$

$$\rightarrow\frac{a^2}{2}\left[\sum_{\text{links}}(E^a_{\vec{p}})^2+\sum_{\text{plaquettes}}(B_{\vec{p}})^2\right]. \tag{2.3}$$

The canonical commutators become on the lattice (Latin superscripts are vector indices)

$$[A^a_{\vec{p}},E^b_{\vec{p}'}]=-i\frac{1}{a^2}\,\delta_{p_x p'_x}\delta_{p_y p'_y}\delta_{ab}. \tag{2.4}$$

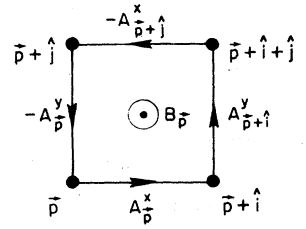

FIG. 2. Identification of the plaquette variable $B_{\vec{p}}$ in terms of link variables $A^a_{\vec{p}}$.

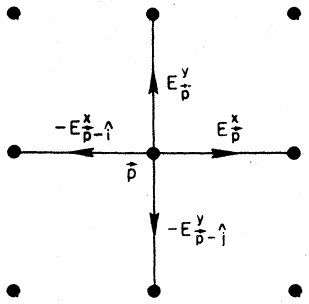

FIG. 3. $(\vec{\nabla}\cdot\vec{E})_{\vec{p}}$ shown in terms of contributing links.

Equations (2.3) and (2.4) define a version of the lattice QED that is closely parallel to the continuum theory and is referred to as the noncompact version since the variable B can assume arbitrarily large values.

Alternative versions of latticized QED replace B^2 in (2.3) by other "potentials" $V(B)$. In particular, the Hamiltonian in the compact formulation developed by Wilson and by Kogut and Susskind is equivalent to

$$H^{(2)} \equiv a^2\left[\tfrac{1}{2}\sum_{\text{links}}(E^a_{\vec{p}})^2 + \frac{1}{e^2a^4}\sum_{\text{plaquettes}}(1-\cos ea^2B_{\vec{p}})\right], \tag{2.5}$$

where the coupling constant e has the dimension of (length)$^{-1/2}$ in two dimensions. Equation (2.5) reduces to (2.3) in the $ea^2 \to 0$ limit.[5] However, (2.5), in contrast to (2.3), depends on the magnitude of the charge, and the higher-order terms in B^2 give rise to nonlinear corrections to a free photon description.

Henceforth all variables will be made dimensionless by dividing through by appropriate powers of a, which will be set equal to unity. Further, we represent the new dimensionless e as g and canonically rescale

$$\begin{aligned} A &\to A/g, \\ B &\to B/g, \\ E &\to gE, \end{aligned} \tag{2.6}$$

so that our two Hamiltonians become

$$H^{(1)} = \tfrac{1}{2}\left(g^2\sum_{\text{links}}E^2 + \frac{1}{g^2}\sum_{\text{plaquettes}}B^2\right), \tag{2.7}$$

$$H^{(2)} = \tfrac{1}{2}\left(g^2\sum_{\text{links}}E^2 + \frac{2}{g^2}\sum_{\text{plaquettes}}(1-\cos B)\right). \tag{2.8}$$

B. The Hilbert space and gauge fixing

We must next specify the Hilbert space on which our operators act. In order to exhibit more clearly the crucial role played by the specific form chosen for the Hamiltonian we depart from earlier formulations of the compact theory[2] and choose to realize the commutation relations (2.4) for both the compact and noncompact versions of QED by interpreting $\vec{A}_{\vec{p}}$ and $\vec{E}_{\vec{p}}$ as the operators of multiplication and differentiation on the space of square integrable functions of the variables $\vec{A}_{\vec{p}}$ $(-\infty \le A^a_{\vec{p}} \le \infty)$.

The next step is to recognize that not all of the $\vec{E}_{\vec{p}}$'s are truly quantum variables. First we notice that there is only one variable $B_{\vec{p}}$ but several variables $E^a_{\vec{p}}$ for each plaquette. Hence we would like to rewrite the kinetic part of the Hamiltonian in terms of the variables conjugate to the $B_{\vec{p}}$'s plus those linear combinations of $E^a_{\vec{p}}$'s which commute with all of the $B_{\vec{p}}$'s and hence can be diagonalized along with H. This is readily achieved if we note that certain linear combinations of the $\vec{E}$'s are the generators of the time-independent gauge transformations, which commute with H. To be specific, in the $A_0 = 0$ gauge, the Hamiltonian is unchanged if we make the transformation

$$A^a_{\vec{p}} \to A^a_{\vec{p}} + (\vec{\nabla}\Lambda)^a_{\vec{p}}, \tag{2.9}$$

where the lattice gradient is defined as

$$\begin{aligned} (\vec{\nabla}\Lambda)^x_{\vec{p}} &= \Lambda_{\vec{p}+\hat{i}} - \Lambda_{\vec{p}}, \\ (\vec{\nabla}\Lambda)^y_{\vec{p}} &= \Lambda_{\vec{p}+\hat{j}} - \Lambda_{\vec{p}}. \end{aligned} \tag{2.10}$$

From (2.4) it follows that this transformation is effected by the operator

$$U(\{\Lambda_{\vec{p}}\}) = \exp\left[i\sum_{\vec{p},a}(\vec{\nabla}\Lambda)^a_{\vec{p}}E^a_{\vec{p}}\right] \equiv \exp\left(-i\sum_{\vec{p}}\Lambda_{\vec{p}}G_{\vec{p}}\right). \tag{2.11}$$

Noting that, for $\Lambda(\infty) = 0$,

$$\sum(\vec{\nabla}\Lambda)^a_{\vec{p}}E^a_{\vec{p}} = -\sum_{\vec{p}}\Lambda_{\vec{p}}(\vec{\nabla}\cdot\vec{E})_{\vec{p}}, \tag{2.12}$$

we can identify the gauge generators G by

$$G_{\vec{p}} = (\vec{\nabla}\cdot\vec{E})_{\vec{p}}. \tag{2.13}$$

Since $U(\{\Lambda_{\vec{p}}\})$ and the generators $G_{\vec{p}}$ commute with H, as well as with all physical observables, we can diagonalize them and work within any individual eigensubspace. The eigenvalues of $G_{\vec{p}}$ may be interpreted as static external charges $\rho_{\vec{p}}$, and in this way we see that the eigenvalue equation is nothing but Gauss's law, i.e.,

$$[(\vec{\nabla}\cdot\vec{E})_{\vec{p}} - \rho_{\vec{p}}]\,|\psi\rangle = 0, \tag{2.14}$$

for all states $|\psi\rangle$ in this sector of our Hilbert space.

Restricting ourselves henceforth to any such eigensubspace we now decompose the electric field into a classical (longitudinal) and a quantum

(transverse) part, writing

$$\vec{E} = \vec{E}^L + \vec{E}^T , \tag{2.15}$$

where $\vec{E}^L$ and $\vec{E}^T$ are defined by the conditions

$$(\vec{\nabla} \times \vec{E}^L) = 0 , \qquad \vec{\nabla} \cdot \vec{E}^T = 0 . \tag{2.16}$$

The lattice curl and divergence in (2.16) are defined as in (2.1) and (2.2). Now, (2.16) implies that

$$\vec{E}^L = -\vec{\nabla}\phi \tag{2.17}$$

and, by (2.14), ϕ satisfies

$$\vec{\nabla} \cdot \vec{E} = \vec{\nabla} \cdot \vec{E}^L = -\nabla^2 \phi = \rho , \tag{2.18}$$

where the lattice Laplacian is defined to be[6]

$$(\nabla^2 \phi)_{\vec{p}} = \phi_{\vec{p}+\hat{i}} + \phi_{\vec{p}-\hat{i}} + \phi_{\vec{p}+\hat{j}} + \phi_{\vec{p}-\hat{j}} - 4\phi_{\vec{p}} . \tag{2.19}$$

We observe that (2.18) and (2.19) are compatible only if $\sum_{\vec{p}} \rho_{\vec{p}} = 0$; however, this constraint is automatically satisfied.[7] Since

$$\sum_{\vec{p}} \phi (\nabla^2 \phi)_{\vec{p}} = -\sum_{\vec{p}} (\vec{\nabla}\phi)_{\vec{p}}^{\,2} \leq 0 ,$$

the homogeneous equation $(\nabla^2 \phi)_{\vec{p}} = 0$ is solved only by a constant ϕ. However, it is clear from (2.17) that any constant in ϕ does not affect the physical variables $\vec{E}^L$. We can therefore restrict ϕ to lie in the space of functions $\sum \phi_{\vec{p}} = 0$, on which (2.18) can be inverted uniquely to give ϕ and hence $\vec{E}^L$.

As for $\vec{E}^T$, (2.16) implies

$$\vec{E}^T = \vec{\nabla} \times L , \tag{2.20}$$

where L is a pseudovector (loop variable) defined on each plaquette of the lattice, and zero outside,[8] and

$$\begin{aligned} (\vec{\nabla} \times L)^x_{\vec{p}} &\equiv L_{\vec{p}} - L_{\vec{p}-\hat{j}} , \\ (\vec{\nabla} \times L)^y_{\vec{p}} &\equiv -L_{\vec{p}} + L_{\vec{p}-\hat{i}} . \end{aligned} \tag{2.21}$$

Equation (2.21) can be inverted to give

$$L_{\vec{p}} = \sum_{j=-\infty}^{p_y} E^{Tx}_{(p_x, j)} . \tag{2.22}$$

Summing up, for each link of the lattice,

$$\begin{aligned} E^x_{\vec{p}} &= -(\phi_{\vec{p}+\hat{i}} - \phi_{\vec{p}}) + (L_{\vec{p}} - L_{\vec{p}-\hat{j}}) , \\ E^y_{\vec{p}} &= -(\phi_{\vec{p}+\hat{j}} - \phi_{\vec{p}}) - (L_{\vec{p}} - L_{\vec{p}-\hat{i}}) . \end{aligned} \tag{2.23}$$

From (2.22) and (2.4) we can easily deduce

$$[L_{\vec{p}}, B_{\vec{p}'}] = i\delta_{p_x p'_x} \delta_{p_y p'_y} . \tag{2.24}$$

Thus L and B are conjugate quantum variables. We now rewrite the Hamiltonian as

$$\begin{aligned} H &= H_{\text{kin}} + H_{\text{pot}} , \\ H_{\text{kin}} &= \frac{g^2}{2} \sum_{\text{links}} E^2 \\ &= \frac{g^2}{2} \sum_{\text{links}} [(E^L)^2 + 2\vec{E}^L \cdot \vec{E}^T + (E^T)^2] . \end{aligned} \tag{2.25}$$

In the continuum,

$$\int dV \vec{E}^L \cdot \vec{E}^T = \int dV \phi \vec{\nabla} \cdot \vec{E}^T = 0$$

and the same integration by parts is easily demonstrated on the lattice. Hence we are left with the usual c-number Coulomb term plus the dynamical term written in terms of variables conjugate to the B's, i.e.,

$$\begin{aligned} H_{\text{kin}} &= \frac{g^2}{2} \sum_{\text{links}} [(E^L)^2 + (E^T)^2] \\ &= \frac{g^2}{2} \sum_{\text{links}} [(\vec{\nabla}\phi)^2 + (\vec{\nabla} \times L)^2] . \end{aligned} \tag{2.26}$$

The potential terms in the two versions are, respectively,

$$H^{(1)}_{\text{pot}} = \frac{1}{2g^2} \sum_{\text{plaquettes}} B^2 \tag{2.27}$$

and

$$H^{(2)}_{\text{pot}} = \frac{1}{g^2} \sum_{\text{plaquettes}} (1 - \cos B) . \tag{2.28}$$

C. Periodicity and the compact Hamiltonian

The periodicity of the potential term (2.28) in the compact version introduces crucial differences between it and the noncompact theory. When we add (2.26) to the potential term (2.27) for the noncompact version we have a two-dimensional array of coupled harmonic oscillators, which is a straightforward problem to solve. In momentum space one finds simply the spectrum of discretized oscillators. However, the potential (2.28) defines a problem that is far from trivial. As a consequence of the periodicity, there exists an infinite set of operators

$$T_{\vec{p}} = e^{2\pi i L_{\vec{p}}} , \tag{2.29}$$

which shift $B_{\vec{p}}$ by 2π and thus commute with the Hamiltonian $H = H^{\text{kin}} + H^{(2)}_{\text{pot}}$ as well as the gauge generators. Furthermore, the $T_{\vec{p}}$'s are unitary operators and hence their eigenvalues are phases which can be written as $e^{2\pi i \epsilon_{\vec{p}}}$, where $-\frac{1}{2} \leq \epsilon_{\vec{p}} \leq \frac{1}{2}$.[9]

The problem of simultaneously diagonalizing the $T_{\vec{p}}$ and H is reminiscent of the Bloch wave problem for conduction electrons in a periodic potential.[10]

The states corresponding to definite eigenvalues $\epsilon_{\vec{p}}$ satisfy

$$\psi(\{B_{\vec{p}}+2\pi n_{\vec{p}}\})=\exp\left(-2\pi i\sum_{\vec{p}} n_{\vec{p}}\epsilon_{\vec{p}}\right)\psi(\{B_{\vec{p}}\}) \quad (2.30)$$

and hence can be expressed as

$$\psi(\{B_{\vec{p}}\})=\exp\left(-i\sum_{\vec{p}}\epsilon_{\vec{p}}B_{\vec{p}}\right)\tilde{\psi}(\{B_{\vec{p}}\}), \quad (2.31)$$

where $\tilde{\psi}(\{B_{\vec{p}}\})$ has period 2π. From this it follows that in all computations we can with no loss of generality restrict $B_{\vec{p}}$ to the range $[-\pi,\pi]$ denoting it by the angular variable $-\theta_{\vec{p}}$. This explains our use of the term compact for this version of the theory. We represent the canonically conjugate operators $L_{\vec{p}}$ as differentiation with respect to $\theta_{\vec{p}}$, that is,

$$L_{\vec{p}}=\frac{1}{i}\frac{\partial}{\partial\theta_{\vec{p}}}. \quad (2.32)$$

Notice that (2.31) tells us that we could equivalently restrict our wave functions to be always periodic functions of the variables $\theta_{\vec{p}}$ and redefine $L_{\vec{p}}$ to be

$$L_{\vec{p}}=\frac{1}{i}\frac{\partial}{\partial\theta_{\vec{p}}}+\epsilon_{\vec{p}}. \quad (2.33)$$

In this representation the Hamiltonian becomes

$$\begin{aligned}H=&\frac{g^2}{2}\sum_{\text{links}}(\vec{\nabla}\phi)^2\\&+\frac{g^2}{2}\sum_{\vec{p}}\left[\left(\frac{1}{i}\frac{\partial}{\partial\theta_{\vec{p}+\hat{j}}}-\frac{1}{i}\frac{\partial}{\partial\theta_{\vec{p}}}+\epsilon_{\vec{p}+\hat{j}}-\epsilon_{\vec{p}}\right)^2\right.\\&\left.+\left(\frac{1}{i}\frac{\partial}{\partial\theta_{\vec{p}+\hat{i}}}-\frac{1}{i}\frac{\partial}{\partial\theta_{\vec{p}}}+\epsilon_{\vec{p}+\hat{i}}-\epsilon_{\vec{p}}\right)^2\right]\\&+\frac{1}{g^2}\sum_{\vec{p}}(1-\cos\theta_{\vec{p}})\end{aligned} \quad (2.34)$$

and operates in a space of periodic functions

$$\tilde{\psi}(\{\theta_{\vec{p}}\})=\tilde{\psi}(\{\theta_{\vec{p}}+2\pi n_{\vec{p}}\}), \quad (2.35)$$

where again all integrations may be restricted to $-\pi\le\theta_{\vec{p}}\le\pi$. In either case (2.32) or (2.33) the spectrum of $L_{\vec{p}}$ is

$$L_{\vec{p}}=m_{\vec{p}}+\epsilon_{\vec{p}},\quad m_{\vec{p}}=0,\pm1,\pm2,\ldots.$$

When $\rho_{\vec{p}}=0$ we identify the physical sector as that in which the time-averaged $\vec{E}$ field is everywhere zero. In other words, along with Wilson, we set $\epsilon_{\vec{p}}=0$ when no charges are present.

D. Introducing charges

To complete our formulation of QED on a lattice we now specify the way in which quantum charges are introduced. Referring to previous studies of fermions on a lattice[11] we write the Hamiltonian for electrons

$$H^{\text{particle}}=M\sum_{\vec{p}}\psi^\dagger_{\vec{p}}\psi_{\vec{p}}+H^{\text{particle}}_{\text{kin}}, \quad (2.36)$$

where $H^{\text{particle}}_{\text{kin}}$ is written as a sum over gauge-invariant operators of the form

$$\psi^\dagger_{\vec{p}}\exp\left(-i\sum_{\vec{p}'=\vec{p}}^{\vec{p}+(D-1)\hat{a}}A^a_{\vec{p}'}\right)\psi_{\vec{p}+D\hat{a}}, \quad (2.37)$$

with suitable weighting factors so that in the $g\to0$ limit $H^{\text{particle}}_{\text{kin}}$ becomes the correct free fermion lattice kinetic energy. The operators (2.37) insert only integer flux on links of the lattice.

The form of H_{kin} given by (2.37) is not the most general expression one can write consistent with the requirement that the Hamiltonian be gauge invariant. As already noted there are many inequivalent Hamiltonian formulations possible. Our goal in this paper is to study confinement (or the lack of it) in two extreme cases: ordinary noncompact QED and a simple generalization of the compact version of the theory formulated by Wilson and Kogut and Susskind. Since in the compact version the eigenvalue of $\vec{E}$ on any link is always an integer we are led to the choice (2.37). As we have previously discussed causality requires that the quantum fluctuations which create a fermion pair are accompanied by changes in the electric fields only in some finite region of space around the pair. There are clearly less restrictive ways of satisfying the dual constraints of gauge invariance and causality than (2.37). We have not studied these in any detail; however, it is obvious that by spreading the flux in the initial state (with fractional flux along each of several paths) one tends to weaken rather than strengthen the confinement. The form (2.37) represents an extreme case, and we find that even in this extreme case confinement does not persist at weak coupling in 3+1 dimensions.

It is clear that each operator of the form specified in (2.37) creates a pair of opposite charges at two separated points joined by a string of unit electric flux created by

$$U_{\vec{p},\vec{p}+D\hat{a}}=\exp\left(-i\sum_{\vec{p}'=\vec{p}}^{\vec{p}+(D-1)\hat{a}}A^a_{\vec{p}'}\right).$$

Formally we show this by observing that if we start from a state defined by

$$E^a_{\vec{p}}|\{\mathcal{E}^a_{\vec{p}}\}\rangle=\mathcal{E}^a_{\vec{p}}|\{\mathcal{E}^a_{\vec{p}}\}\rangle \quad (2.38)$$

we find, using the commutator (2.4),

$$E^a_{\vec{p}} U_{\vec{p}',\vec{p}'+D\hat{\delta}} |\{\mathcal{S}^a_{\vec{\rho}}\}\rangle = [E^a_{\vec{p}}, U_{\vec{p}',\vec{p}'+D\hat{\delta}}] |\{\mathcal{S}^a_{\vec{\rho}}\}\rangle + \mathcal{S}^a_{\vec{p}} U |\{\mathcal{S}^a_{\vec{\rho}}\}\rangle$$

$$= \left(\mathcal{S}^a_{\vec{p}} + \delta_{ab} \sum_{r=0}^{D-1} \delta_{\vec{p},\vec{p}'+r\hat{\delta}}\right) U |\{\mathcal{S}^a_{\vec{\rho}}\}\rangle . \quad (2.39)$$

It is also clear from the structure of the fermion operators in (2.37) that only states of zero total charge, containing equal numbers of particles and antiparticles, are created from the vacuum by the Hamiltonian. Now the general problem of interest is to compute the ground-state energy in the sector of the Hilbert space containing such a configuration. In the static, or large M, limit we can find the eigenstates of (2.34) plus (2.36) by studying pair states created by applying operators of the form (2.37) to the ground state of the theory (2.34).

Since the $U_{\vec{p},\vec{p}+D\hat{a}}$ *do not* commute with the operators $T_{\vec{p}}$ (2.29) the eigenvalues $\epsilon_{\vec{p}}$ are changed when a state with pairs is created by the operation (2.37). The specific change in values of $\epsilon_{\vec{p}}$ can be computed directly from Gauss's law as follows. $U_{\vec{p},\vec{p}+\hat{a}}$ creates a string of unit field strength along the link from $\vec{p}$ to $\vec{p}+\hat{a}$ and thereby changes $(\vec{\nabla}\cdot\vec{E})$ by $+1$ unit at $\vec{p}$ and by -1 unit at $\vec{p}+\hat{a}$. By (2.18) this means a change in the static Coulomb field owing to the additional field of a dipole pair with $+1$ unit of charge at $\vec{p}$ and -1 unit at $\vec{p}+\hat{a}$. According to (2.23) and (2.33) there must then be a compensating change in $\epsilon_{\vec{p}}$ defined by $L \to L + \delta\epsilon \equiv ULU^{-1}$ such that

$$(\vec{\nabla} \times \delta\epsilon)_{\vec{p}} = \vec{E}^{\text{string}}_{\vec{p}} - \vec{E}^{\text{Coul}}_{\vec{p}} . \quad (2.40)$$

Since the static Coulomb dipole field, $\vec{E}^{\text{Coul}}_{\vec{p}}$, in general has fractional units of flux on each link, it follows that the $\delta\epsilon_{\vec{p}}$ are nonzero. In particular if we start with a state with no charges and $\epsilon_{\vec{p}} \equiv 0$, there will necessarily be nonvanishing (fractional) $\epsilon_{\vec{p}}$ everywhere in the sector of states with charges present.

III. THE ONE-PLAQUETTE UNIVERSE

We turn now to the problem of a very tiny universe made of a single square, or plaquette, as illustrated in Fig. 4. Our reason for doing so is that this very simple problem allows us to present most of our calculational techniques as well as much of the physics of the more interesting problem of a lattice whose linear dimension, $(2N_0+1)a$, is arbitrarily large.

Let us examine first the noncompact Hamiltonian [(2.26)+(2.27)] which for a single square becomes

$$H^{(1)} = \frac{g^2}{2} \sum_{\text{links}} (\vec{\nabla}\phi)^2 + \frac{g^2}{2}(4L^2) + \frac{1}{2g^2} B^2 , \quad (3.1)$$

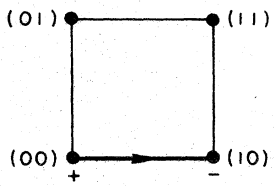

FIG. 4. A single-plaquette universe showing the charge configuration and notation discussed in the text.

i.e., each of the four links contributes $g^2[\frac{1}{2}L^2 + \frac{1}{2}(\vec{\nabla}\phi)^2]$ to the energy. It is clear that for this Hamiltonian the value of the parameter g appears in the classical Coulomb energy as determined from (2.18), but is entirely irrelevant for the transverse dynamical part from which it can be rescaled away by undoing (2.6). Furthermore this is a trivial theory to solve. The energy of the ground state for any charge distribution is the energy of the Coulomb configuration corresponding to that distribution plus the ground-state energy of a harmonic oscillator of frequency $\omega = 2$. Thus the difference in energy between a state with no charges and a state with charge $+1$ at $(0,0)$ and charge -1 at $(1,0)$ as illustrated in Fig. 4 is clearly just the Coulomb energy of that state.

It is also interesting to compute the expectation values of the electric field strengths created on the links of the plaquette by the presence of the dipole in Fig. 4. Let us denote by

$$|\rho_{00} = \rho_{01} = \rho_{10} = \rho_{11} = 0; \psi_0\rangle$$

the state in the sector of no charges with the oscillator in (3.1) in the ground state. Obviously for this sector

$$\vec{E}^{\text{Coul}} = 0 , \quad \langle \rho_{\vec{p}} = 0; \psi_0 | L | \rho_{\vec{p}} = 0; \psi_0 \rangle = 0 ,$$

so that the expectation value of $E^a_{\vec{p}}$ vanishes.

Let us now consider the expectation value of the electric field components E^x_{00}, E^y_{10}, E^x_{01}, and E^y_{00} in the state

$$e^{-iA^x_{00}} |\rho_{\vec{p}} = 0; \psi_0\rangle$$
$$= |\rho_{00} = 1, \rho_{10} = -1, \rho_{01} = \rho_{11} = 0; \psi\rangle , \quad (3.2)$$

with the charge dipole present. It follows from (2.17) and (2.18) that the Coulomb field corresponding to this charge distribution is

$$E^{\text{Coul},x}_{00} = \tfrac{3}{4} ,$$
$$E^{\text{Coul},y}_{00} = E^{\text{Coul},x}_{01} = -E^{\text{Coul},y}_{10} = \tfrac{1}{4} . \quad (3.3)$$

Since $e^{-iA^x_{00}}$ creates a unit string $E^x_{00} = 1$, (2.23) and (3.3) tell us that $\langle\psi|L|\psi\rangle = \frac{1}{4}$ at $t = 0$. Since by (3.1) this state describes a coherent oscillator with $\omega = 2$ the time-dependent expectation values of the elec-

tric field are

$$\langle E^x_{00}(t)\rangle = \tfrac{3}{4} + \tfrac{1}{4}\cos\omega t\,,$$
$$\langle E^y_{00}(t)\rangle = \langle E^x_{01}(t)\rangle = -\langle E^y_{10}(t)\rangle = \tfrac{1}{4} - \tfrac{1}{4}\cos\omega t\,. \tag{3.4}$$

This state describes a static Coulomb configuration plus an oscillating photon cloud. Even though the cloud oscillates it is clear that the time-averaged value of the $\vec{E}$ field in this state is exactly the Coulomb value. The oscillating nature of the cloud is an artifact of our very small "universe"—the radiation cloud cannot radiate away because it hits the nearby boundaries of this small system and is reflected back. In an infinite system the coherent cloud would simply radiate away, unshielding the Coulomb field of the two charges.

We now examine the same problem in the compact version:

$$H^{(2)} = \mathfrak{E}_{\text{Coul}} + \frac{g^2}{2}\,4L^2 + \frac{1}{g^2}(1-\cos B)\,, \tag{3.5}$$

where $\mathfrak{E}_{\text{Coul}}$ is the Coulomb energy of the charge configuration. Since B and L are conjugate variables, $e^{\pm iB}$ are simply raising and lowering operators, viz.,

$$[L, e^{\pm iB}] = \mp e^{\pm iB} \tag{3.6}$$

and so we can write in an L basis

$$H^{(2)} = \mathfrak{E}_{\text{Coul}} + \frac{g^2}{2}\,4(m+\epsilon)^2 + \frac{1}{2g^2}(2 - J^+ - J^-)\,, \tag{3.7}$$

where $\epsilon = 0$ for the configuration with no charges and $\epsilon = \frac{1}{4}$ with the Coulomb configuration (3.3), using (2.40), $J^{\pm} \equiv e^{\mp iB}$. Alternatively we can work in a B basis, treating L as the momentum. For this we introduce

$$-\theta = B\,,$$
$$L = \frac{1}{i}\frac{\partial}{\partial\theta} + \epsilon = \frac{1}{i}\frac{\partial}{\partial\theta} + \tfrac{1}{4}\,, \tag{3.8}$$

requiring that the eigenfunctions of $(1/i)\partial/\partial\theta$ be periodic on the interval $-\pi \leq \theta \leq \pi$ as is necessary for integer eigenvalues. Since the problem in this representation is to solve the Hamiltonian

$$H^{(2)} = \mathfrak{E}_{\text{Coul}} + 2g^2\left(\frac{1}{i}\frac{\partial}{\partial\theta} + \epsilon\right)^2 + \frac{1}{g^2}(1-\cos\theta) \tag{3.9}$$

in the space of periodic functions on the interval $[-\pi, \pi]$, it is a precise analog of the Bloch wave problem for a Schrödinger particle in a periodic potential.[10]

For strong coupling, $g^2 \gg 1$, (3.7) is dominated by the momentum term $4(g^2/2)(m+\epsilon)^2$. Since $-\frac{1}{2} \leq \epsilon \leq \frac{1}{2}$, the ground state is clearly $m = 0$ and has energy $E = \mathfrak{E}_{\text{Coul}} + 2g^2\epsilon^2$. We notice that for $g \to \infty$ this is an energy eigenstate. In the no charge sector $\rho_{\vec{p}} = 0$, $\epsilon = 0$ the $\vec{E}$ field vanishes identically, whereas in the sector with the pair of charges as in Fig. 4, $\epsilon = \frac{1}{4}$ and $E^x_{00} = 1$, $E^y_{00} = E^x_{01} = E^y_{10} = 0$. In this strong-coupling limit the "photons" do not radiate away but remain to focus the electric field on the string between the charges. This result continues to apply for a larger lattice.

For weak coupling, $g^2 \ll 1$, we can use the arguments usually given to obtain an approximate solution to the corresponding Bloch wave problem. The ground-state energy has contributions owing to barrier penetration.[10] Keeping only the leading correction due to tunneling between neighboring minima of the periodic potential we have

$$E(\epsilon) = E^{\text{harmonic}} - A(g^2)e^{-B/g^2}\cos 2\pi\epsilon\,.$$

These tunneling corrections, although very small, are nonanalytic in g^2 at $g^2 = 0$ and depend upon the charges through the ϵ distribution. They are the crucial new feature in the compact theory. We now give a variational estimate of these terms and discuss physical effects.

The terms in (3.9) proportional to ϵ are removed in terms of the Bloch momentum by introducing a trial wave function of the form [undoing (2.31) and (2.33)]

$$\psi = e^{-i\epsilon\theta}\chi(\theta)\,,$$
$$\chi(\theta + 2\pi) = e^{+2\pi i\epsilon}\chi(\theta)\,. \tag{3.10}$$

Since the potential term in (3.9) is a very deep well, a reasonable trial form for the variational calculation is a narrow Gaussian packet[12] centered about $\theta = 0$. However, in order to satisfy the boundary conditions (3.10) we must construct a superposition of the form

$$\chi(\theta) = \sum_{n=-\infty}^{\infty} e^{2\pi i\epsilon n}\psi_{\text{tr}}(\theta - 2\pi n)\,, \tag{3.11}$$

where $\psi_{\text{tr}}(\theta - 2\pi n)$ is the oscillator ground-state trial wave function for the potential centered at $\theta = 2\pi n$. The ground-state energy is then just the energy of the ground state for a single such well plus correction terms which correspond to the overlap between two wells with different values of n. For $g^2 \ll 1$ the dominant correction comes of course from tunneling between nearest wells. To illustrate our variational procedure we write the trial function

$$\psi_{\text{tr}}(\theta) = e^{-(\gamma/2)\theta^2}\,, \tag{3.12}$$

where γ is the variational parameter which we expect to be large for small g^2. The normalization integral is then

$$\langle\chi|\chi\rangle=\int_{-\pi}^{\pi}d\theta\,\chi^*(\theta)\chi(\theta)=\sum_{n_1,n_2=-\infty}^{\infty}\int_{-\pi}^{\pi}d\theta\exp[2\pi i\epsilon(n_1-n_2)]\exp[-\tfrac{1}{2}\gamma(\theta-2\pi n_1)^2]\exp[-\tfrac{1}{2}\gamma(\theta-2\pi n_2)^2]. \tag{3.13}$$

One of the n sums can be done, since it simply extends the range of integration to $(-\infty,+\infty)$. To wit, if we redefine variables

$$N=n_1-n_2,\quad n=n_1,$$

(3.13) becomes

$$\langle\chi|\chi\rangle=\sum_{N=-\infty}^{\infty}e^{2\pi i\epsilon N}\int_{-\infty}^{\infty}d\theta\exp(-\tfrac{1}{2}\gamma\theta^2)\exp[-\tfrac{1}{2}\gamma(\theta+2\pi N)^2]\,. \tag{3.14}$$

Therefore we have for each N a simple Gaussian integral. The dominant contribution, for narrow packets, comes from $N=0$ and the leading corrections are $\sim e^{-\gamma\pi^2}\cos 2\pi\epsilon\ll 1$ corresponding to tunneling between neighboring wells with $N=\pm 1$. The variational energy is

$$\begin{aligned}E(\gamma)=\langle H\rangle&=\langle\chi|H|\chi\rangle/\langle\chi|\chi\rangle\\&=\mathcal{E}_{\text{Coul}}+\left(g^2\gamma+\frac{1-e^{-1/4\gamma}}{g^2}\right)+\frac{2\sum_{N>0}e^{-\gamma\pi^2N^2}\cos(2\pi\epsilon N)\{-2g^2\gamma^2\pi^2N^2+(1/g^2)e^{-1/4\gamma}[1-(-1)^N]\}}{1+2\sum_{N>0}e^{-\gamma\pi^2N^2}\cos(2\pi\epsilon N)}\end{aligned} \tag{3.15}$$

and γ is fixed by minimizing $E(\gamma)$.[12] For $g^2\ll 1$, the term in (3.15) which is independent of ϵ and N dominates and gives $\gamma\approx 1/2g^2\gg 1$. We find in this way, up to higher-order tunneling corrections,

$$\begin{aligned}E(\gamma)-\mathcal{E}_{\text{Coul}}&=1+(O(g^2),\ \epsilon\text{-independent terms})\\&\quad-\frac{2}{g^2}e^{-\pi^2/2g^2}\cos 2\pi\epsilon\left[\frac{\pi^2-4}{2}+O(g^2)\right]+O(e^{-\pi^2/g^2}).\end{aligned} \tag{3.16}$$

The first term on the right-hand side is, to leading order in g, the zero-point energy for an oscillator of frequency 2 and mass $1/4g^2$, just as in the noncompact theory. The remaining terms are the tunneling corrections in the compact formulation. Since they depend on ϵ they introduce, in addition to the Coulomb energy, a contribution to the energy which depends on the positions of the interacting charges. This contribution is nonanalytic in the charge at $g=0$ and is the crucial new term in the compact theory which is responsible for confinement on a large lattice. The role this plays can be understood better if we compute the time-averaged value of the electric field strength. This average can be computed directly since, from (3.9),

$$\langle E^a_{\vec{p}}\rangle=-(\vec{\nabla}\phi)^a_{\vec{p}}+\frac{1}{4g^2}\frac{\partial}{\partial\epsilon}\langle H\rangle. \tag{3.17}$$

By (3.16) we have

$$\begin{aligned}\langle E^x_{00}\rangle&=\tfrac{3}{4}+\frac{\pi}{g^4}\left(\frac{\pi^2-4}{2}\right)e^{-\pi^2/2g^2}\sin 2\pi\epsilon\,,\\\langle E^y_{00}\rangle&=\langle E^x_{01}\rangle=-\langle E^y_{10}\rangle\\&=\tfrac{1}{4}-\frac{\pi}{g^4}\left(\frac{\pi^2-4}{2}\right)e^{-\pi^2/2g^2}\sin 2\pi\epsilon\,.\end{aligned} \tag{3.18}$$

We see that in addition to the Coulomb component there is a residual effect from the coherent cloud which is proportional to $(1/g^4)e^{-\pi^2/2g^2}$. The sign of this added tunneling contribution, with $\epsilon=\frac{1}{4}$ (corresponding to the charge configuration in Fig. 4), is such as to increase the strength of the field along the link between the charges, while at the same time decreasing its other components relative to their Coulomb values (3.3). Thus its effect is to focus the field along the link joining the dipole. In the next section our study of the large lattice in 2+1 dimensions also reveals such a focusing. This leads, for large separations, to an energy which is proportional to the distance between the charges. This is the dynamical origin of the linear confinement in (2+1)-dimensional QED first described for $g^2\ll 1$ by Polyakov.

The trial function (3.12) is also a good variational guess for the strong-coupling region where $g^2\gg 1$ and $\gamma\to 0$. We then have to sum contributions from large values of N to the variational energy since all $N^2\lesssim(\pi^2\gamma)^{-1}$ will contribute significantly to (3.15). This can be done by transforming to a dual form of the periodic Gaussian using the Poisson sum formula

$$\sum_{n=-\infty}^{\infty}g(n)=\sum_{m=-\infty}^{\infty}\int_{-\infty}^{\infty}d\phi\,e^{2\pi im\phi}g(\phi). \tag{3.19}$$

In (3.11) this gives

$$\chi(\theta)=\sum_{m=-\infty}^{\infty}\int_{-\infty}^{\infty}d\phi\, e^{2\pi i m\phi}e^{2\pi i\phi\epsilon}e^{-(\gamma/2)(\theta-2\pi\phi)^2}$$

$$=\frac{1}{\sqrt{2\pi\gamma}}e^{i\epsilon\theta}\sum_{m=-\infty}^{\infty}e^{im\theta}e^{-(m+\epsilon)^2/2\gamma^2}, \qquad (3.20)$$

which approaches

$$\chi(\theta)\propto e^{i\epsilon\theta},$$

for $g\to\infty$, which is the region $\gamma\ll 1$. Recalling (3.10) we have for the ground-state trial solution

$$\psi=\text{constant},$$

which is the exact ground-state solution for (3.7) in the strong-coupling limit $g\to\infty$ ($m=0$). Since it does well in both the strong- and weak-coupling limits, (3.12) is presumably a reasonable trial form for studying intermediate coupling as well.[13]

IV. TWO-DIMENSIONAL ABELIAN THEORY

We now turn to the compact formulation of the two-dimensional Abelian theory on a lattice of $(2N_0+1)\times(2N_0+1)$ points. As discussed in Sec. II, our interest is in calculating the energy of a pair of oppositely charged particles as a function of their separation, for the coupling region $g^2\ll 1$. Under the assumption that the configuration with no charges corresponds to the sector $\epsilon_{\vec{\mathfrak{p}}}=0$, the problem is to compute the ground-state energy of the theory in the sector defined by the $\epsilon_{\vec{\mathfrak{p}}}$ determined from (2.39), i.e.,

$$(\vec{\nabla}\times\epsilon)_{\vec{\mathfrak{p}}}=\vec{E}^{\text{string}}_{\vec{\mathfrak{p}}}-\vec{E}^{\text{Coul}}_{\vec{\mathfrak{p}}}. \qquad (4.1)$$

Here $\vec{E}^{\text{string}}_{\vec{\mathfrak{p}}}$ is unity for links lying on the line between the charges at $(-D/2,0)$ and $(D/2,0)$, (see Fig. 5) and is zero otherwise, and $\vec{E}^{\text{Coul}}_{\vec{\mathfrak{p}}}$ is defined

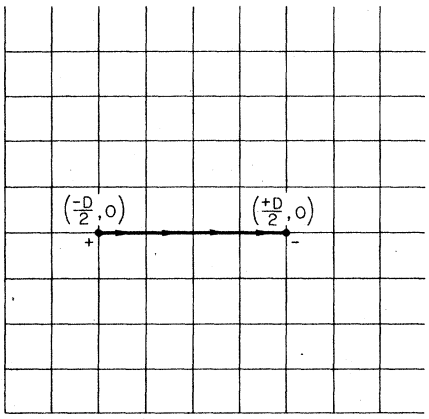

FIG. 5. The larger two-dimensional lattice showing the charge configuration of interest, and the $\vec{E}$-field string of the strong-coupling ground state. Periodic boundary conditions are indicated by extra links at the edges.

by

$$\vec{E}^{\text{Coul}}_{\vec{\mathfrak{p}}}=-(\vec{\nabla}\phi)_{\vec{\mathfrak{p}}}, \qquad (4.2)$$

with

$$-\nabla^2\phi_{\vec{\mathfrak{p}}}=(\delta_{p_x,-D/2}-\delta_{p_x,D/2})\delta_{p_y,0}. \qquad (4.3)$$

Once we have determined the $\epsilon_{\vec{\mathfrak{p}}}$ distribution for this dipole configuration, our problem is to solve for the ground state of the Hamiltonian (2.34) and (2.35).

A. The strong-coupling limit

When $g^2\gg 1$ the kinetic term

$$H_{\text{kin}}=\mathfrak{E}_{\text{Coul}}+\frac{g^2}{2}\sum_{\text{links}}(\vec{\nabla}\times L)^2, \qquad (4.4)$$

dominates the compact Hamiltonian since the potential is a bounded operator; therefore in the $g^2\to\infty$ limit eigenstates of H are products of eigenstates of $L_{\vec{\mathfrak{p}}}$, i.e.,

$$L_{\vec{\mathfrak{p}}}\,|\{m_{\vec{\mathfrak{p}}}\}\rangle=(m_{\vec{\mathfrak{p}}}+\epsilon_{\vec{\mathfrak{p}}})\,|\{m_{\vec{\mathfrak{p}}}\}\rangle \qquad (4.5)$$

[see (2.33)]. If we start with the state specified by the $m_{\vec{\mathfrak{p}}}$ determined in (4.1) (i.e., all $m_{\vec{\mathfrak{p}}}=0$ for this case of the dipole),[14] we find for the energy in the $g^2\to\infty$ limit

$$\langle H\rangle_{g^2\to\infty}=\mathfrak{E}_{\text{Coul}}+\frac{g^2}{2}\sum_{\text{links}}(\vec{\nabla}\times\epsilon)^2$$

$$=\frac{g^2}{2}\sum_{\text{links}}[(-\vec{\nabla}\phi)^2+(\vec{\nabla}\times\epsilon)^2]$$

$$=\frac{g^2}{2}\sum_{\text{links}}(-\vec{\nabla}\phi+\vec{\nabla}\times\epsilon)^2. \qquad (4.6)$$

Referring back to (4.1) and (4.2) we have

$$\langle H\rangle=\frac{g^2}{2}\sum_{\text{links}}(E^{\text{string}})^2=\frac{g^2}{2}D, \qquad (4.7)$$

where D is the length of the string. We can also shift the m's away from zero in order to construct other states in this ϵ sector. In such states the original string is lengthened by the creation of new string segments and/or loops according to (4.5). The energy of these states will thus be higher than that given in (4.7) for the state with all $m=0$ which is hence the ground state.[15] In short, in the $g^2\to\infty$ limit the coherent cloud of transverse photons corresponding to the eigenstate $|m_{\vec{\mathfrak{p}}}=0\rangle$ holds itself together for all time and focuses the Coulomb field to a string, giving rise to a linear confining potential in (4.6).

B. The weak-coupling limit

The weak-coupling analysis is more complicated and, as we saw for one plaquette, essential con-

tributions of the photon self-interactions will be missed in a perturbative expansion about $g^2=0$. The procedure we will use can be justified by the recursive variational technique which we have used in other applications.[1] Fortunately, however, very little of the physics we wish to discuss in the $g^2 \ll 1$ limit depends upon a detailed knowledge of how this calculation is carried out, and we can *finesse* these complications by using information gleaned from our analysis of the one-plaquette problem. In that case we saw that in (3.15) the $N=0$ terms dominate the expectation value of the energy, which is sensitive only to very small values of $\theta \approx g \ll 1$. Those terms corresponding to $N \neq 0$, which arise from our prescription rendering the wave function periodic, were sensitive to values of $|\theta| \approx \pi$, and hence were reduced by factors of $e^{-\pi^2/g^2} \ll 1$. Therefore for the purpose of determining the form of the trial wave function we could both forget that the variable is restricted to $|\theta| \leq \pi$ and replace $(1-\cos\theta)/g^2$ by $(1/2g^2)\theta^2$.

Returning now to the Hamiltonian on the two-dimensional lattice we first change representation as we did for one plaquette so that the operator $(1/i)\partial/\partial\theta_{\vec{p}}$ has eigenvalues $(m_{\vec{p}}+\epsilon_{\vec{p}})$. In this case, (2.34) becomes

$$H=\mathfrak{E}_{\text{Coul}}-\frac{g^2}{2}\sum_{\vec{p}}\left(\frac{\partial}{\partial\theta_{\vec{p}}}-\frac{\partial}{\partial\theta_{\vec{p}+\hat{j}}}\right)^2 -\frac{g^2}{2}\sum_{\vec{p}}\left(\frac{\partial}{\partial\theta_{\vec{p}}}-\frac{\partial}{\partial\theta_{\vec{p}+\hat{i}}}\right)^2+V(\{\theta_{\vec{p}}\}) \tag{4.8}$$

and we must require that the eigenfunctions satisfy

$$\psi(\{\theta_{\vec{p}}+2\pi n_{\vec{p}}\})=\exp\left(2\pi i\sum_{\vec{p}}\epsilon_{\vec{p}}n_{\vec{p}}\right)\psi(\{\theta_{\vec{p}}\}). \tag{4.9}$$

[In (4.8) we can define the contribution of links at the edge of the lattice by $L_{\vec{p}}=0$ if either p_x or p_y is less than $-N$ or greater than or equal to N. Alternatively we can impose periodic boundary conditions by defining

$$E^x_{N_0,p_y}=E^x_{-N_0-1,p_y}, \quad E^y_{p_x,N_0}=E^y_{p_x,-N_0-1}.$$

The difference between these two formulations is only of order $1/V$, and for convenience we shall use the periodic formulation in our subsequent discussion.]

Having imposed the proper aperiodicity by (4.9) we can replace the potential in (4.8) by its quadratic approximation for the purpose of determining the form of the ground-state trial wave function. This approximation turns (4.8) into a massless free field theory and the ground-state wave function of this system is easily written in momentum space as a product of Gaussians in the variables

$$\tilde{\theta}_{\vec{k}}=\frac{1}{\sqrt{V}}\sum_{\vec{p}}e^{i\vec{k}\cdot\vec{p}}\theta_{\vec{p}}, \quad V=(2N_0+1)^2. \tag{4.10}$$

Incorporating the boundary conditions (4.9) this choice gives a trial wave function of the form

$$\psi_{\text{trial}}(\{\theta_{\vec{p}}\})=\prod_{\vec{p}}\left(\sum_{n_{\vec{p}}=-\infty}^{\infty}\right)\exp\left(2\pi i\sum_{\vec{p}}n_{\vec{p}}\epsilon_{\vec{p}}\right)\psi_0(\{\theta_{\vec{p}}-2\pi n_{\vec{p}}\}),$$

with

$$\psi_0(\{\theta_{\vec{p}}\})=\exp\left(-\frac{1}{2g^2}\sum_{\vec{p},\vec{p}'}\theta_{\vec{p}}\Delta_{\vec{p}\vec{p}'}\theta_{\vec{p}'}\right)\delta\left(\sum_{\vec{p}}\theta_{\vec{p}}\right), \tag{4.11}$$

where

$$\Delta_{\vec{p}\vec{p}'}=\frac{1}{V}\sum_{\vec{k}}e^{i\vec{k}\cdot(\vec{p}-\vec{p}')}\gamma_{\vec{k}} \tag{4.12}$$

and

$$\vec{k}=\frac{\pi\vec{n}}{N_0}, \quad -N_0\leq n_x,n_y\leq N_0.$$

The δ function on the sum of the angles that appears in (4.11), viz., $\delta(\sum_{\vec{p}}\theta_{\vec{p}})$, arises from the fact that the kinetic term in the Hamiltonian (4.8) contains only differences of the canonical momenta[8] and hence $\sum_{\vec{p}}\theta_{\vec{p}}$, or by (4.10), $\tilde{\theta}_{\vec{k}=0}$ is a classical variable that we can fix at an arbitrary constant value. In general we can treat $\gamma_{\vec{k}}$ as a variational function to be determined for arbitrary g^2 using the recursive variational procedures developed earlier. In the weak-coupling limit of replacing the potential by its quadratic approximation, i.e.,

$$V(\{\theta_{\vec{p}}\})\to\frac{1}{2g^2}\sum_{\vec{p}}\theta_{\vec{p}}^2, \quad g^2\ll 1, \tag{4.13}$$

the Hamiltonian reduces to non-self-interacting photons on a lattice and we can solve for the $\gamma_{\vec{k}}$ directly:

$$\gamma_{\vec{k}}=(4-2\cos k_x-2\cos k_y)^{-1/2}, \quad \vec{k}\neq 0. \tag{4.14}$$

However, owing to the aperiodicity conditions there are nonperturbative contributions to the energy even in the weak-coupling limit as in the one-plaquette analysis of the preceding section.

We turn next to a calculation of these contributions, using (4.11) and (4.12) as our trial functions, to

$$E(\gamma)=\langle\psi_{\text{trial}}|H|\psi_{\text{trial}}\rangle/\langle\psi_{\text{trial}}|\psi_{\text{trial}}\rangle. \tag{4.15}$$

Straightforward generalization of the manipulations used for the one-square problem yield

$$E(\gamma)=\mathcal{E}_{\text{Coul}}+\tfrac{1}{4}\sum_{\vec{k}}\gamma_{\vec{k}}(4-2\cos k_x-2\cos k_y)+\frac{1}{g^2}\sum_{\vec{p}}\langle 1-\cos\theta_{\vec{p}}\rangle_0$$

$$-\frac{\pi^2}{2g^2}\sum_{\vec{k}}\gamma_{\vec{k}}^2(4-2\cos k_x-2\cos k_y)\langle\tilde{N}_{\vec{k}}\tilde{N}_{-\vec{k}}\rangle+\frac{1}{g^2}\sum_{\vec{p}}\langle 1-(-1)^{N_{\vec{p}}}\rangle\langle\cos\theta_{\vec{p}}\rangle_0\,, \tag{4.16}$$

where we have used the same trick in carrying out one of the $n_{\vec{p}}$ sums for each $\vec{p}$ as in (3.14), and we define

$$\langle f(\theta_{\vec{q}})\rangle_0=\frac{\prod_{\vec{p}}\left(\int_{-\infty}^{\infty}d\theta_{\vec{p}}\right)\exp\left[-(1/g^2)\sum_{\vec{p}\vec{p}'}\theta_{\vec{p}}\Delta_{\vec{p}\vec{p}'}\theta_{\vec{p}'}\right]f(\theta_{\vec{q}})\,\delta\left(\sum_{\vec{p}}\theta_{\vec{p}}\right)}{\prod_{\vec{p}}\left(\int_{-\infty}^{\infty}d\theta_{\vec{p}}\right)\exp\left[-(1/g^2)\sum_{\vec{p}\vec{p}'}\theta_{\vec{p}}\Delta_{\vec{p}\vec{p}'}\theta_{\vec{p}'}\right]\delta\left(\sum_{\vec{p}}\theta_{\vec{p}}\right)} \tag{4.17}$$

and

$$\langle f(N_{\vec{q}})\rangle=\frac{\prod_{\vec{p}}\left(\sum_{N_{\vec{p}}=-\infty}^{\infty}\right)\exp\left[-(\pi^2/g^2)\sum_{\vec{p}\vec{p}'}N_{\vec{p}}\Delta_{\vec{p}\vec{p}'}N_{\vec{p}'}+2\pi i\sum_{\vec{p}}N_{\vec{p}}\epsilon_{\vec{p}}\right]f(N_{\vec{q}})\,\delta\left(\sum_{\vec{p}}N_{\vec{p}}\right)}{\prod_{\vec{p}}\left(\sum_{N_{\vec{p}}=-\infty}^{\infty}\right)\exp\left[-(\pi^2/g^2)\sum_{\vec{p}\vec{p}'}N_{\vec{p}}\Delta_{\vec{p}\vec{p}'}N_{\vec{p}'}+2\pi i\sum_{\vec{p}}N_{\vec{p}}\epsilon_{\vec{p}}\right]\delta\left(\sum_{\vec{p}}N_{\vec{p}}\right)}\,, \tag{4.18}$$

with

$$\tilde{N}_{\vec{k}}=\frac{1}{\sqrt{V}}\sum_{\vec{p}}e^{i\vec{k}\cdot\vec{p}}N_{\vec{p}}\,. \tag{4.19}$$

The constraint on the N-sums $\delta(\sum_{\vec{p}}N_{\vec{p}})$ arises from the constraint $\delta(\sum\theta_{\vec{p}})$ in (4.11) and tells us that the overlap of initial and final states vanishes except for values of $N_{\vec{p}}$ such that $\sum_{\vec{p}}N_{\vec{p}}=0$.

It is clear from (4.16)–(4.18) that the $N_{\vec{p}}=0$ terms dominate the energy for weak coupling, as they did in the one-plaquette example, since the $N_{\vec{p}}\neq 0$ contributions are suppressed by tunneling factors $\sim e^{-(\text{const})/g^2}$. A very good approximation to $\gamma_{\vec{k}}$ is obtained if we ignore the $N_{\vec{p}}\neq 0$ terms in (4.16) and variationally compute $\partial E/\partial\gamma_{\vec{k}}$ for all $\vec{k}\neq 0$. This gives (4.14) up to corrections of order g^2 due to the difference between the potential term (2.28) and its quadratic approximation (4.13). Substituting into (4.16) and using (4.17) and (4.19) leads to

$$E(\{\epsilon\})=\mathcal{E}_{\text{Coul}}+\tfrac{1}{2}\sum_{\vec{k}}(4-2\cos k_x-2\cos k_y)^{1/2}$$

$$+(O(g^2),\ \epsilon\ \text{independent})$$

$$-\frac{1}{2g^2}\sum_{\vec{p}}\langle\pi^2N_{\vec{p}}{}^2-2[1-(-1)^{N_{\vec{p}}}]\langle\cos\theta_{\vec{p}}\rangle_0\rangle\,, \tag{4.20}$$

where

$$\langle\cos\theta_{\vec{p}}\rangle_0=\exp\left(-\frac{g^2}{4V}\sum_{\vec{k}}\gamma_{\vec{k}}^{-1}\right)=1+O(g^2)\,. \tag{4.21}$$

This ends the quantum-mechanical part of our problem; all that remains is to do the sums over $N_{\vec{p}}$ as defined by (4.18). The rest of this section will be devoted to doing these sums.

C. Why doing the $N_{\vec{p}}$ sums by brute force does not work

Calculation of the sums over $N_{\vec{p}}$ in (4.20) can be reduced to the evaluation of a single normalization factor $Z(\{\epsilon\})$ since from (4.18)

$$\langle f(N_{\vec{q}})\rangle=\frac{1}{Z(\{\epsilon\})}f\left(\frac{1}{2\pi i}\frac{\partial}{\partial\epsilon_{\vec{q}}}\right)Z(\{\epsilon\})\,, \tag{4.22}$$

where

$$Z(\{\epsilon\})=\prod_{\vec{p}}\left(\sum_{N_{\vec{p}}=-\infty}^{\infty}\right)\exp\left(-\frac{\pi^2}{g^2}\sum_{\vec{p}\vec{p}'}N_{\vec{p}}\Delta_{\vec{p}\vec{p}'}N_{\vec{p}'}+2\pi i\sum_{\vec{p}}N_{\vec{p}}\epsilon_{\vec{p}}\right)\delta\left(\sum_{\vec{p}}N_{\vec{p}}\right), \tag{4.23}$$

$Z(\{\epsilon\})$ can be thought of as the partition function[16] for a neutral gas of charged particles which interact via the potential $\Delta_{\vec{p}\vec{p}'}$. For such a system one expects that the free-energy density, defined by

$$F(\{\epsilon\})=\frac{-1}{V}\ln Z(\{\epsilon\})\,, \tag{4.24}$$

is well defined in the infinite-volume limit. An expansion of Z in powers of F will give terms of order $V^mF^m/m!$. Since there are $O(V^m)$ terms with m nonvanishing $N_{\vec{p}}$ in the N sum (4.23) one is tempted to try to resum by identifying such terms with the terms of mth and lower order in F. This resummation can be performed if the interaction has sufficiently short range: This is the Mayer cluster expansion in statistical mechanics.[16]

Let us illustrate this, for $\epsilon_{\vec{p}}=0$, for the extreme case $\Delta_{\vec{p}\vec{p}'}=0$ for all $\vec{p}\neq\vec{p}'$. We then can readily see that

$$1 \leq \prod_{\vec{p}} \left(\sum_{N_{\vec{p}}=-\infty}^{\infty} \right) \delta\left(\sum N_{\vec{p}} \right) \exp\left(-\frac{\pi^2}{g^2} \Delta_0 \sum_{\vec{p}} N_{\vec{p}}^2 \right)$$

$$\equiv e^{-VF_0} \leq \prod_{p \neq p_0} \left[\sum_{N_{\vec{p}}=-\infty}^{\infty} \exp\left(-\frac{\pi^2}{g^2} \Delta_0 N_{\vec{p}}^2 \right) \right], \quad (4.25)$$

with

$$\Delta_0 = \Delta_{\vec{p}\vec{p}}.$$

The sum on the right-hand side of (4.25) is finite and we see that

$$0 \geq F_0 \geq -\ln\left[1 + 2 \sum_{N=1}^{\infty} \exp\left(-\frac{\pi^2}{g^2} \Delta_0 N^2 \right) \right], \quad (4.26)$$

which shows that F_0 is well defined for $V \to \infty$.

We now examine the change in this result when we include the correct $\Delta_{\vec{p}\vec{p}'}$ for $\vec{p} \neq \vec{p}'$, considering the case $\epsilon_{\vec{p}} \equiv 0$. We rewrite (4.23) in the form

$$Z(\{0\}) = \prod_{\vec{p}} \left(\sum_{N_{\vec{p}}=-\infty}^{\infty} \right) \delta\left(\sum N_{\vec{p}} \right) \times \exp\left(-\frac{\pi^2}{g^2} \Delta_0 \sum_{\vec{p}} N_{\vec{p}}^2 \right) \prod_{\vec{p} \neq \vec{p}'} (1 + f_{\vec{p}\vec{p}'}), \quad (4.27)$$

where

$$f_{\vec{p}\vec{p}'} = \exp\left(-\frac{\pi^2}{g^2} N_{\vec{p}} \Delta_{\vec{p}\vec{p}'} N_{\vec{p}'} \right) - 1. \quad (4.28)$$

We can attempt to evaluate the correction to F_0 from the nonvanishing $f_{\vec{p}\vec{p}'}$ by keeping terms with successively higher numbers of f's. In this approximation the first correction to VF_0 is of the form of $\sum_{\vec{p} \neq \vec{p}'} f_{\vec{p}\vec{p}'}$, which must grow no faster than V if this expansion procedure is to make any sense. However, $\Delta_{\vec{p}\vec{p}'}$, hence $f_{\vec{p}\vec{p}'}$, fall off only as fast as $1/|\vec{p} - \vec{p}'|$ for large separations $|\vec{p} - \vec{p}'|$. Hence $\sum_{\vec{p} \neq \vec{p}'} f_{\vec{p}\vec{p}'}$ is proportional to $V^{3/2}$, and this method of evaluating the N sums does not work. The problem evidently is a consequence of the long range nature of $\Delta_{\vec{p}\vec{p}'}$. We will now discuss a procedure for resumming (4.23) in a way which avoids these volume divergence difficulties.

D. Doing the $N_{\vec{p}}$ sums by Feynman graphs

In order to develop a correct resummation procedure we rewrite $Z(\{\epsilon_{\vec{p}}\})$ by making use of the identity[17]

$$f(\{N_{\vec{p}}\}) = \prod_{\vec{p}} \left[\int_{-\infty}^{\infty} d\phi_{\vec{p}} \int d\eta_{\vec{p}} e^{2\pi i \phi_{\vec{p}} (\eta_{\vec{p}} - N_{\vec{p}})} \right] f(\{\eta_{\vec{p}}\}). \quad (4.29)$$

Substituting (4.29) in (4.23) and doing the $\eta_{\vec{p}}$ integrations yields, up to irrelevant normalization factors,

$$Z(\{\epsilon_{\vec{p}}\}) \propto \prod_{\vec{p}} \left(\sum_{N_{\vec{p}}=-\infty}^{\infty} \int_{-\infty}^{\infty} d\phi_{\vec{p}} \right) \exp\left(2\pi i \sum_{\vec{p}} N_{\vec{p}} \phi_{\vec{p}} \right) \times \exp\left(-g^2 \sum_{\vec{p}\vec{p}'} (\phi_{\vec{p}} + \epsilon_{\vec{p}}) \Delta^{-1}_{\vec{p}\vec{p}'} (\phi_{\vec{p}'} + \epsilon_{\vec{p}'}) \right), \quad (4.30)$$

where

$$\Delta^{-1}_{\vec{p}\vec{p}'} = \frac{1}{V} \sum_{\vec{k}} e^{i\vec{k} \cdot (\vec{p} - \vec{p}')} \gamma_{\vec{k}}^{-1}. \quad (4.31)$$

Note that the exclusion of the $\vec{k} = 0$ mode is irrelevant in $\Delta^{-1}_{\vec{p}\vec{p}'}$ since $\gamma_{\vec{k}}^{-1} = (4 - 2\cos k_x - 2\cos k_y)$ vanishes for $k_x = k_y = 0$, and so the constraint $\delta(\sum_{\vec{p}} N_{\vec{p}})$ plays no role in the evaluation of (4.30).

One can now rewrite $Z(\{\epsilon\})$ as $Z = Z_1(\{\epsilon\})$ with

$$Z_\lambda(\{\epsilon\}) = \prod_{\vec{p}} \left[\int_{-\infty}^{\infty} d\phi_{\vec{p}} \left(1 + 2\lambda \sum_{N_{\vec{p}}=1}^{\infty} \cos 2\pi N_{\vec{p}} \phi_{\vec{p}} \right) \right] \times \exp\left(-g^2 \sum_{\vec{p}\vec{p}'} (\phi + \epsilon)_{\vec{p}} \Delta^{-1}_{\vec{p}\vec{p}'} (\phi + \epsilon)_{\vec{p}'} \right). \quad (4.32)$$

The parameter λ has been introduced in (4.32) in order to simplify subsequent bookkeeping. The term of order λ^m in (4.32) is clearly identical to the term in the original N sum, (4.23), corresponding to configurations with m nonvanishing $N_{\vec{p}}$. Hence, an expansion in powers of λ has the same volume divergence diseases as the N sums. However, the reformulation of the λ sum in (4.32) has an important advantage in that it allows us to relate the expansion in powers of λ to a summation of Feynman graphs. In this way we can convert the problem of volume divergences to that of infrared divergences of Feynman graphs and use well-known techniques for resumming the series in λ so as to avoid all problems.

First we simplify the calculation by truncating the N sums in (4.32) to $N_{\vec{p}} = 1$. This can be justified by evaluating also the sums keeping $N_{\vec{p}} = 1, 2, 3, \ldots, N_{\max}$ and showing that the additional contributions to $\langle N_{\vec{p}}^2 \rangle$ and $\langle 1 - (-1)^{N_{\vec{p}}} \rangle$ are higher-order corrections to our result.[18] We also temporarily set all $\epsilon_{\vec{p}} = 0$ in (4.32) since our resolution of the volume divergence problem can be demonstrated in this simple case. We will of course reinstate the proper values of $\epsilon_{\vec{p}}$ in order to calculate the energy.

The relation of the λ series in (4.32) to a sum of Green's functions which can be evaluated by Feynman graph techniques can be made explicit if we observe that

$$\psi(\{\phi_{\vec{p}}\}) = \exp\left(-\frac{g^2}{2}\sum_{\vec{p}\vec{p}'}\phi_{\vec{p}}\Delta^{-1}_{\vec{p}\vec{p}'}\phi_{\vec{p}'}\right) \tag{4.33}$$

is the ground-state wave function of a free massless field theory whose propagator is

$$\frac{1}{g^2}\Delta_{\vec{p}\vec{p}'} = \frac{1}{g^2 V}\sum_{\vec{k}} e^{i\vec{k}\cdot(\vec{p}-\vec{p}')}\frac{1}{\omega_{\vec{k}}}, \qquad \omega_{\vec{k}} = \gamma_{\vec{k}}^{-1} = (4 - 2\cos k_x - 2\cos k_y)^{1/2}. \tag{4.34}$$

The Feynman graphs for the first few terms in the expansion of $\langle\cos(2\pi\phi_{\vec{p}})\rangle$ and $\langle\cos(2\pi\phi_{\vec{p}})\cos(2\pi\phi_{\vec{p}'})\rangle$ are shown in Figs. 6 and 7, respectively. Graphs such as these are valuable guides in resumming appropriate subsets of terms so as to avoid all volume divergence problems. The decomposition into connected and disconnected terms shown in Fig. 7 yields the exponentiation desired in (4.24) if the connected graphs grow no faster than V for $V\to\infty$. Figure 8(a) shows a typical term in the calculation of the order λ^2 term $\langle\cos(2\pi\phi_{\vec{p}})\cos(2\pi\phi_{\vec{p}'})\rangle^{\text{connected}}$ and it is obvious that all remaining loop integrations are infrared divergent. However, Fig. 8(b) shows that in every higher order of λ there are graphs corresponding to insertions of tadpoles anywhere on the propagators joining points $\vec{p}$ and $\vec{p}'$. The sum over all such tadpoles on each line corresponds to modifying the propagator, viz.,

$$-\frac{1}{g^2\omega_{\vec{k}}} \to \frac{-1}{g^2\omega_{\vec{k}}} + \frac{-1}{g^2\omega_{\vec{k}}}\mu^2\frac{-1}{g^2\omega_{\vec{k}}} + \frac{-1}{g^2\omega_{\vec{k}}}\mu^2\frac{-1}{g^2\omega_{\vec{k}}}\mu^2\frac{-1}{g^2\omega_{\vec{k}}} + \cdots \equiv \frac{-1}{g^2\tilde{\omega}_{\vec{k}}(\mu^2)}, \tag{4.35}$$

where

$$\tilde{\omega}_{\vec{k}}(\mu^2) = \omega_{\vec{k}} + \frac{\mu^2}{g^2} \quad\text{and}\quad \mu^2 \propto \langle\cos 2\pi\phi_{\vec{p}}\rangle.$$

The new propagator $\tilde{\omega}_{\vec{k}}(\mu^2)$ is no longer singular at $k_x = k_y = 0$ and correspondingly the disease preventing Mayer clustering in the preceding section is cured by this procedure. It is clear, however, that no finite approximation to the series (4.35) can cure the disease.

In practice manipulating pieces of the graphical summation in this way is cumbersome. We can, however, obtain the same result efficiently, now

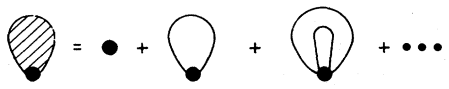

FIG. 6. Contributions to $\langle\cos 2\pi\phi_{\vec{p}}\rangle$. Each loop corresponds to a sum over spatial momenta $(1/V)\sum_{\vec{k}}{}_{V\to\infty}(1/4\pi^2)\int d^2k$ and each line is a factor $-1/g\omega_{\vec{k}}$.

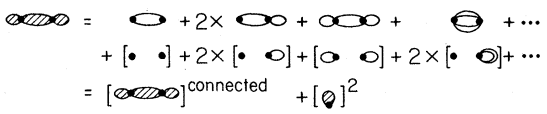

FIG. 7. Contributions to $\langle\cos 2\pi\phi_{\vec{p}}\cos 2\pi\phi_{\vec{p}'}\rangle$.

that we recognize what we are looking for. We rewrite, still keeping $\epsilon = 0$,

$$Z_1(\{0\}) = \prod_{\vec{p}}\left[\int_{-\infty}^{\infty} d\phi_{\vec{p}}[(1+2\cos 2\pi\phi_{\vec{p}})e^{+\mu^2\phi_{\vec{p}}^2}]\right] \times \exp\left[-\left(\sum_{\vec{p}\vec{p}'} g^2\phi_{\vec{p}}\Delta^{-1}_{\vec{p}\vec{p}'}\phi_{\vec{p}'} + \mu^2\sum_{\vec{p}}\phi_{\vec{p}}^2\right)\right]. \tag{4.36}$$

In (4.36) μ^2 is defined so that when one normal orders the bracket $[(1+2\cos 2\pi\phi_{\vec{p}})e^{\mu^2\phi_{\vec{p}}^2}]$ with respect to the new propagator $(1/g^2)[\Delta^{-1}_{\vec{p}\vec{p}'} + (\mu^2/g^2)\delta_{\vec{p}\vec{p}'}]^{-1}$ then the coefficient of $:\phi^2:$ vanishes. Specifically we define a_m by

$$(1+2\cos 2\pi\phi_{\vec{p}})e^{\mu^2\phi_{\vec{p}}^2} = \sum_{m=0}^{\infty} a_m :\phi_{\vec{p}}^{2m}: \tag{4.37}$$

and choose μ^2 so that $a_1 \equiv 0$.

Equation (4.37) simplifies considerably for $g^2 \ll 1$, in which case, self-consistently, μ^2 is very small so that we can expand to first order in μ^2. This gives

$$\mu^2 = 4\pi^2\exp\left[-\frac{\pi^2}{g^2 V}\sum_{\vec{k}}\left(\frac{1}{\omega_{\vec{k}}}\right)\right], \tag{4.38}$$

with higher-order contributions exponentially damped by powers of e^{-1/g^2}.

The essential accomplishment of the resummation indicated in (4.36)–(4.38) is to construct a propagator in $Z(\{\epsilon\})$ that has sufficiently strong screening to guarantee that a loopwise expansion of the graphs contributing to $Z_\lambda(\{\epsilon\})$ will be infra-

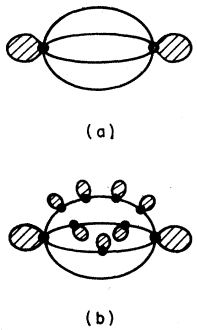

FIG. 8. (a) Diagram of order λ^2 contributing to a connected two-point function. (b) Diagram of order λ^9 which corresponds to mass insertions in Fig. 8(a).

red finite. In fact the new propagator

$$\frac{1}{g^2}\Delta_{\vec{p}\vec{p}'}{}^{\mu}=\frac{1}{g^2V}\sum_{\vec{k}}\frac{e^{i\vec{k}\cdot(\vec{p}-\vec{p}')}}{\omega_{\vec{k}}+\mu^2/g^2}\propto\frac{g^2}{\mu^4|\vec{p}-\vec{p}'|^3}\ , \tag{4.39}$$

for a separation $|\vec{p}-\vec{p}'|$ greater than the shielding length g^2/μ^2. We can now finally turn to the evaluation of $E(\{\epsilon\})$ in (4.20).

E. Evaluation of $E(\{\epsilon\})$

With the restriction of $N_{\vec{p}}$ to $0,\pm1$, (4.20) and (4.22) simplify to

$$E(\{\epsilon_{\vec{p}}\})=\mathcal{E}_{\text{Coul}}+\tfrac{1}{2}\sum_{\vec{k}}\omega_{\vec{k}}+(O(g^2),\epsilon\text{ independent})+\frac{1}{8\pi^2g^2}\frac{1}{Z(\{\epsilon\})}\sum_{\vec{p}}(\pi^2-4\langle\cos\theta_{\vec{p}}\rangle_0)\frac{\partial^2Z(\{\epsilon\})}{\partial\epsilon_{\vec{p}}{}^2}\ ,$$

with

$$Z(\{\epsilon_{\vec{p}}\})=\prod_{\vec{p}}\left[\int_{-\infty}^{\infty}d\phi_{\vec{p}}[(1+2\cos2\pi\phi_{\vec{p}})e^{\mu^2\phi_{\vec{p}}{}^2}]\right]\exp\left[-\left(\sum_{\vec{p}\vec{p}'}g^2(\phi_{\vec{p}}+\epsilon_{\vec{p}})\Delta^{-1}{}_{\vec{p}\vec{p}'}(\phi_{\vec{p}'}+\epsilon_{\vec{p}'})+\mu^2\sum_{\vec{p}}\phi_{\vec{p}}{}^2\right)\right] \tag{4.40}$$

and we can, to leading order in $g^2\ll1$, replace $\langle\cos(\theta_{\vec{p}})\rangle_0$ by unity in accord with (4.21). By the definition of μ^2, the bracket contains no terms in $:\phi_{\vec{p}}{}^2:$. To leading order in e^{-1/g^2} the bracket can now be replaced by unity since the $:\phi_{\vec{p}}{}^4:$ and higher-power terms are all reduced by powers of $\mu^2\sim e^{-1/g^2}$.[19]

In this approximation we find that

$$Z(\{\epsilon_{\vec{p}}\})=\exp\left[-\mu^2\sum_{\vec{p}_1\vec{p}_2\vec{p}_3}\epsilon_{\vec{p}_1}(g^2\Delta^{-1}+\mu^2)^{-1}{}_{\vec{p}_1\vec{p}_2}(g^2\Delta^{-1})_{\vec{p}_2\vec{p}_3}\epsilon_{\vec{p}_3}\right], \tag{4.41}$$

which implies that the energy changes from its $\epsilon=0$ value by

$$\delta E(\{\epsilon\})=\frac{\pi^2-4}{2\pi^2g^2}\mu^4\sum_{\vec{p}_1\vec{p}_2\vec{p}_3}\epsilon_{\vec{p}_1}(g^2\Delta^{-1}+\mu^2)^{-2}{}_{\vec{p}_1\vec{p}_2}(g^2\Delta^{-1})^2{}_{\vec{p}_2\vec{p}_3}\epsilon_{\vec{p}_3}\ . \tag{4.42}$$

Rewriting this in k space,

$$\delta E(\{\epsilon\})=\frac{\pi^2-4}{2\pi^2g^2}\mu^4\sum_{\vec{p}\vec{p}'}\epsilon_{\vec{p}}\left[\frac{1}{V}\sum_{\vec{k}}e^{i\vec{k}\cdot(\vec{p}-\vec{p}')}\frac{\omega_{\vec{k}}^2}{(\omega_{\vec{k}}+\mu^2/g^2)^2}\right]\epsilon_{\vec{p}'}\ . \tag{4.43}$$

Inserting $\omega_{\vec{k}}=4-2\cos k_x-2\cos k_y$ and rearranging the sums gives

$$\delta E(\{\epsilon\})=-\frac{(\pi^2-4)}{2\pi^2g^2}\mu^4\sum_{\vec{p}\vec{p}'}\epsilon_{\vec{p}}\frac{1}{V}\sum_{\vec{k}}e^{i\vec{k}\cdot(\vec{p}-\vec{p}')}\left(\omega_{\vec{k}}+\frac{\mu^2}{g^2}\right)^{-2}(\nabla^2\epsilon)_{\vec{p}'}\ , \tag{4.44}$$

where ∇^2 denotes the Laplacian on the lattice as given by (2.19). The $\epsilon_{\vec{p}}$ are defined such that

$$\nabla^2\epsilon_{\vec{p}}=-\vec{\nabla}\times(\vec{\nabla}\times\vec{\epsilon})_{\vec{p}}=-\vec{\nabla}\times(\vec{E}^{\text{string}}-\vec{E}^{\text{Coul}})_{\vec{p}}$$
$$=-\vec{\nabla}\times\vec{E}^{\text{string}}_{\vec{p}}. \tag{4.45}$$

We find, by inserting (4.45) in (4.44) and again "integrating by parts" (rearranging the $\sum_{\vec{p}\vec{p}'}$), that the energy shift is

$$\delta E(\{\epsilon\})=\frac{\pi^2-4}{2\pi^2g^2}\mu^4\sum_{\vec{p}\vec{p}'}(\vec{E}^{\text{string}}_{\vec{p}}-\vec{E}^{\text{Coul}}_{\vec{p}})\cdot\vec{E}^{\text{string}}_{\vec{p}'}I_{\vec{p}\vec{p}'}$$
$$=(\delta E)_{ss}-(\delta E)_{sc}\ , \tag{4.46}$$

where

$$I_{\vec{p}\vec{p}'}=\frac{1}{V}\sum_{\vec{k}}e^{i\vec{k}\cdot(\vec{p}-\vec{p}')}\left[(4-2\cos k_x-2\cos k_y)^{1/2}+\frac{\mu^2}{g^2}\right]^{-2} \tag{4.47}$$

and "ss" means "string-string" energy and "sc" is "string-Coulomb." The contribution in (4.46) proportional to the square of the string field increases linearly with the distance between the charges for large separations $D\gtrsim g^2/\mu^2$. This is shown directly using (4.47) and going to the limit $V\to\infty$,

$$\delta E_{ss}=\frac{(\pi^2-4)}{2\pi^2g^2}\mu^4\int_{-\pi}^{\pi}dk_xdk_y\frac{\sin^2(k_xD)/2}{k_x{}^2[(k_x{}^2+k_y{}^2)^{1/2}+\mu^2/g^2]^2}$$
$$\approx\mu^2D\frac{(\pi^2-4)}{2\pi^4}\int_0^{\pi D}dx\frac{\sin^2x}{x^2}$$
$$\approx\frac{\pi^2-4}{4\pi^3}\mu^2D\quad\text{for}\quad D>\frac{g^2}{\mu^2}\ . \tag{4.48}$$

The contribution from $(\delta E)_{sc}$ is a negligible correction proportional to the Coulomb energy.

Although our methods are very different we have arrived at the same conclusion for QED in 2+1 di-

mensions as Polyakov—namely, that there is linear confinement in the weak-coupling region and hence no "phase transition" encountered as the coupling constant is turned down from strong coupling. The effects of monopoles in Polyakov's description are replaced in the present calculation by the tunneling effects as expressed by the distribution of $N_{\vec{p}}$ in (4.23). The nonzero $N_{\vec{p}}$ have a small but finite density on the lattice according to our calculation of (4.23).

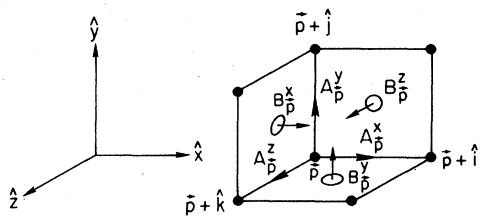

FIG. 9. Labeling of site, link, and plaquette variables for a three-dimensional lattice.

V. THREE SPATIAL DIMENSIONS

Much of the formalism and discussion of the previous sections can readily be extended to a three-dimensional spatial lattice, but there are, as we shall see, crucial differences. Once again we begin with $A^a_{\vec{p}}$ and $E^a_{\vec{p}}$ defined on the link leaving the point $\vec{p}$ in the direction $\hat{a}$. We associate three faces with each point $\vec{p}$ as shown in Fig. 9. As before for each face we can define one gauge-invariant combination of A's,

$$\vec{B}_{\vec{p}} \equiv (\vec{\nabla} \times \vec{A}_{\vec{p}}) ,$$

e.g.,

$$B^x_{\vec{p}} = A^z_{\vec{p}+\hat{j}} - A^z_{\vec{p}} - (A^y_{\vec{p}+\hat{k}} - A^y_{\vec{p}}) , \tag{5.1}$$

etc. A new feature of the three-dimensional lattice is that for every cube there is one redundant variable defined in this way since it follows from (5.1) that

$$(\vec{\nabla} \cdot \vec{B})_{\vec{p}} \equiv B^x_{\vec{p}+\hat{i}} - B^x_{\vec{p}} + B^y_{\vec{p}+\hat{j}} - B^y_{\vec{p}} + B^z_{\vec{p}+\hat{k}} - B^z_{\vec{p}} = 0 . \tag{5.2}$$

Equation (5.2) states that the sum of the outward-pointing B's on the six faces of each elementary cube vanishes, being the lattice version of the divergence of a curl.

We again rewrite the $\vec{E}$ fields, following (2.23), in terms of a Coulomb part plus a transverse field,

$$\begin{aligned} &\vec{E}_{\vec{p}} = -(\vec{\nabla}\phi)_{\vec{p}} + (\vec{\nabla} \times \vec{L})_{\vec{p}} , \\ &-\nabla^2 \phi_{\vec{p}} = \rho_{\vec{p}} . \end{aligned} \tag{5.3}$$

Since $\vec{E}_\rho$ and hence H does not depend on $(\vec{\nabla} \cdot \vec{L})_{\vec{p}}$, $(\vec{\nabla} \cdot \vec{B})_{\vec{p}}$ is a classical time-independent variable and the constraints (5.2) must also be imposed on our trial wave functions at each lattice cube. There was no parallel condition for the two-dimensional theory, in which only the background B, or $\sum_{\vec{p}} \theta_{\vec{p}}$ on the entire lattice, was constrained. Aside from this restriction we can closely follow the calculational techniques discussed previously.

Once again for the compact case we have a potential which is periodic,

$$V = \frac{1}{g^2} \sum_{\text{faces}} (1 - \cos B^a_{\vec{p}}) , \tag{5.4}$$

so that the Hamiltonian is invariant under translations $B \to B + 2\pi n$. Hence, for every face, we can again define a conjugate pair of variables $\{m^a_{\vec{p}}, \theta^a_{\vec{p}}\}$, with $-\pi \leq \theta^a_{\vec{p}} \leq \pi$, and

$$\vec{L}_{\vec{p}} \equiv \vec{\epsilon}_{\vec{p}} + \vec{m}_{\vec{p}} , \quad -\tfrac{1}{2} \leq \epsilon^a_{\vec{p}} \leq \tfrac{1}{2} . \tag{5.5}$$

The Hamiltonian of the quantum field, (4.8) generalized to three dimensions, does not change values of the $\vec{\epsilon}_{\vec{p}}$. Our problem is, as before, to find the ground-state energy of the system with an ϵ distribution obtained by adding a line of unit flux joining two charges separated by a distance D, relative to a configuration with no charges and with all $\vec{\epsilon}_{\vec{p}} = 0$.

The analysis of the strong-coupling limit is not essentially different from the two-dimensional theory and leads to linear confinement as found in (4.7).[20] Turning to the weak-coupling regime of $g^2 \ll 1$, we begin as before by choosing a properly aperiodic trial wave function,

$$\psi_{\text{trial}}(\{\theta^a_{\vec{p}}\}) = \sum_{\{n^a_{\vec{p}}\}} \exp\left(2\pi i \sum_{\vec{p}} \vec{n}_{\vec{p}} \cdot \vec{\epsilon}_{\vec{p}}\right) \exp\left(-\frac{1}{2g^2} \sum_{\vec{p},\vec{p}'} (\vec{\theta}_{\vec{p}} - 2\pi \vec{n}_{\vec{p}}) \cdot \vec{\Delta}_{\vec{p}\vec{p}'} \cdot (\vec{\theta}_{\vec{p}'} - 2\pi \vec{n}_{\vec{p}'})\right) \prod_{\vec{p}} \delta([\vec{\nabla} \cdot (\vec{\theta} - 2\pi \vec{n})]_{\vec{p}}) , \tag{5.6}$$

where the important new constraint on the local divergence of θ has been incorporated in the δ functions. As a result of these δ-function constraints $\vec{\Delta}_{\vec{p}\vec{p}'}$ has the tensor form

$$\Delta^{ab}_{\vec{p}\vec{p}'} = \Delta^{ab}(\vec{p} - \vec{p}') = \frac{1}{V} \sum_{\vec{k}} e^{i\vec{k} \cdot (\vec{p} - \vec{p}')} \gamma_{\vec{k}} \delta_{ab} , \tag{5.7}$$

since the contribution from terms proportional to $k^a k^b$ automatically vanishes. Again the coefficients $\gamma_{\vec{k}}$ are to be determined variationally.

Evaluating the expectation value of the Hamiltonian using this trial wave function we find, analogously to (4.16),

$$\langle E\rangle = \langle\psi_{\text{trial}}|H|\psi_{\text{trial}}\rangle / \langle\psi_{\text{trial}}|\psi_{\text{trial}}\rangle$$
$$= \mathcal{E}_{\text{Coul}} + \tfrac{1}{4}\sum_{\vec{k}} (6 - 2\cos k_x - 2\cos k_y - 2\cos k_z)\gamma_{\vec{k}} + \frac{1}{g^2}\sum_{\text{faces}} \langle 1 - \cos\theta^a_{\vec{p}}\rangle_0 + \Delta E(\epsilon)\,, \tag{5.8}$$

where

$$\langle 1 - \cos\theta^c_{\vec{q}}\rangle_0 \equiv \frac{\prod_{\vec{p},a}\left(\int_{-\infty}^{\infty} d\theta^a_{\vec{p}}\right)\exp\left[-(1/g^2)\sum_{\vec{p},\vec{p}'}\vec{\theta}_{\vec{p}}\cdot\vec{\Delta}_{\vec{p}\vec{p}'}\cdot\vec{\theta}_{\vec{p}'}\right](1-\cos\theta^c_{\vec{q}})\prod_{\vec{p}}\delta((\vec{\nabla}\cdot\vec{\theta})_{\vec{p}})}{\prod_{\vec{p},a}\left(\int_{-\infty}^{\infty} d\theta^a_{\vec{p}}\right)\exp\left[-(1/g^2)\sum_{\vec{p},\vec{p}'}\vec{\theta}_{\vec{p}}\cdot\vec{\Delta}_{\vec{p}\vec{p}'}\cdot\vec{\theta}_{\vec{p}'}\right]\prod_{\vec{p}}\delta((\vec{\nabla}\cdot\vec{\theta})_{\vec{p}})}$$

and

$$\Delta E(\epsilon) \equiv -\frac{\pi^2}{2g^2}\sum_{\vec{k},a}(6 - 2\cos k_x - 2\cos k_y - 2\cos k_z)\,\gamma_{\vec{k}}{}^2\langle\tilde{N}^a_{\vec{k}}\tilde{N}^a_{-\vec{k}}\rangle + \frac{1}{g^2}\sum_{\vec{p},a}\langle[1-(-1)^{N^a_{\vec{p}}}]\rangle\langle\cos\theta^a_{\vec{p}}\rangle_0\,,$$

with

$$\langle f(N^c_{\vec{q}})\rangle \equiv \frac{\sum_{\{N^a_{\vec{p}}\}}\prod_{\vec{p}}\delta((\vec{\nabla}\cdot\vec{N})_{\vec{p}})\exp\left(2\pi i\sum_{\vec{p}}\vec{N}_{\vec{p}}\cdot\vec{\epsilon}_{\vec{p}}\right)\exp\left[-(\pi^2/g^2)\sum_{\vec{p},\vec{p}'}\vec{N}_{\vec{p}}\cdot\vec{\Delta}_{\vec{p}\vec{p}'}\cdot\vec{N}_{\vec{p}'}\right]f(N^c_{\vec{q}})}{\sum_{\{N^a_{\vec{p}}\}}\prod_{\vec{p}}\delta((\vec{\nabla}\cdot\vec{N})_{\vec{p}})\exp\left(2\pi i\sum_{\vec{p}}\vec{N}_{\vec{p}}\cdot\vec{\epsilon}_{\vec{p}}\right)\exp\left[-(\pi^2/g^2)\sum_{\vec{p},\vec{p}'}\vec{N}_{\vec{p}}\cdot\vec{\Delta}_{\vec{p}\vec{p}'}\cdot\vec{N}_{\vec{p}'}\right]}\,. \tag{5.9}$$

As before, for small g^2, it is clear that the energy is dominated by the ϵ-independent terms in (5.8) and hence a good approximation is obtained by choosing $\gamma_{\vec{k}}$ to minimize these terms. This gives to leading order

$$\gamma_{\vec{k}} = \frac{1}{(6 - 2\cos k_x - 2\cos k_y - 2\cos k_z)^{1/2}}\,. \tag{5.10}$$

Once again, as in (4.20), we have reduced the problem to the evaluation of the sums over the N's. However, in three dimensions this turns out to be a somewhat more straightforward procedure than it was in two dimensions. One important difference in this case is the constraints

$$(\vec{\nabla}\cdot\vec{N}_{\vec{p}}) = 0 \tag{5.11}$$

requiring zero divergence of $N^a_{\vec{p}}$ through the faces of each individual cube of the lattice. Evidently, this condition implies that the only allowed configurations are those for which the $\vec{N}$ vectors form loops. The simplest such configuration is shown in Fig. 10. It involves four nonvanishing $N^a_{\vec{p}}$'s. All such configurations can be equally well described in terms of integer variables $\vec{t}_{\vec{p}}$ associated with the links of the lattice, with the definition

$$\vec{N}_{\vec{p}} = \vec{\nabla}\times\vec{t}_{\vec{p}}\,. \tag{5.12}$$

Equation (5.12) gives $\vec{N}_{\vec{p}}$ uniquely for each $\vec{t}_{\vec{p}}$, though clearly the $\vec{t}_{\vec{p}}$ associated with any given $\vec{N}_{\vec{p}}$ are determined only up to an arbitrary gradient. Using (5.12) we can rewrite the phase factor in (5.9)

$$\sum_{\vec{p}}\vec{N}_{\vec{p}}\cdot\vec{\epsilon}_{\vec{p}} = \sum_{\vec{p}}(\vec{\nabla}\times\vec{t}_{\vec{p}})\cdot\vec{\epsilon}_{\vec{p}}$$
$$= -\sum_{\vec{p}}\vec{t}_{\vec{p}}\cdot(\vec{E}^{\text{string}} - \vec{E}^{\text{Coul}})_{\vec{p}}\,. \tag{5.13}$$

Since $\vec{E}^{\text{string}}_{\vec{p}}$ is integer valued, as is $\vec{t}_{\rho}$, the phase factor simplifies to

$$\exp\left(2\pi i\sum_{\vec{p}}\vec{N}_{\vec{p}}\cdot\vec{\epsilon}_{\vec{p}}\right) = \exp\left(2\pi i\sum_{\vec{p}}\vec{t}_{\vec{p}}\cdot\vec{E}^{\text{Coul}}_{\vec{p}}\right).$$

Equation (5.9) can now be rewritten in terms of restricted sums over $\vec{t}_{\vec{p}}$

$$\langle f(N^a_{\vec{q}})\rangle = \frac{\sum'_{\{\vec{t}_{\vec{p}}\}}\exp\left(2\pi i\sum_{\vec{p}}\vec{t}_{\vec{p}}\cdot\vec{E}^{\text{Coul}}_{\vec{p}}\right)\exp\left[-(\pi^2/g^2)\sum_{\vec{p},\vec{p}'}(\vec{\nabla}\times\vec{t})_{\vec{p}}\cdot\vec{\Delta}_{\vec{p}\vec{p}'}\cdot(\vec{\nabla}\times\vec{t})_{\vec{p}'}\right]f((\vec{\nabla}\times\vec{t})^a_{\vec{q}})}{\sum'_{\{\vec{t}_{\vec{p}}\}}\exp\left(2\pi i\sum_{\vec{p}}\vec{t}_{\vec{p}}\cdot\vec{E}^{\text{Coul}}_{\vec{p}}\right)\exp\left[-(\pi^2/g^2)\sum_{\vec{p},\vec{p}'}(\vec{\nabla}\times\vec{t})_{\vec{p}}\cdot\vec{\Delta}_{\vec{p}\vec{p}'}\cdot(\vec{\nabla}\times\vec{t})_{\vec{p}'}\right]} \tag{5.14}$$

where the restricted sum $\sum'_{\{\vec{t}_{\vec{p}}\}}$ means that each configuration of $\{(\vec{\nabla}\times\vec{t})_{\vec{p}}\}$ is included only once.[21] Provided the summation over $\vec{t}_{\vec{p}}$ variables converges we can evaluate this expression as

$$\langle f(N^a_{\vec{q}})\rangle = \frac{1}{Z}f\left(\frac{1}{2\pi i}\left(\vec{\nabla}\times\frac{\partial}{\partial\vec{E}^{\text{Coul}}}\right)^a_{\vec{q}}\right)Z(\{\vec{E}^{\text{Coul}}_{\vec{p}}\})\,, \tag{5.15}$$

where $Z(\{\vec{E}^{\text{Coul}}_{\vec{p}}\})$ is the denominator of (5.14). In the strong-coupling limit, $g \to \infty$, the contribution of large values of $t_{\vec{p}}$ is not suppressed and one should not expect the summation to converge. However, for small g the question of convergence is quite different. We can rewrite the weight factors in (5.14) as

$$\exp\left(-\frac{\pi^2}{g^2} t^i_{\vec{p}} V^{ij}_{\vec{p}\vec{p}'} t^j_{\vec{p}}\right),$$

where

$$V^{ij}_{\vec{p}\vec{p}'} = \frac{1}{V} \sum_{\vec{k}} e^{i\vec{k}\cdot(\vec{p}-\vec{p}')}(k^2\delta^{ij} - k^i k^j)\gamma_{\vec{k}}. \tag{5.16}$$

As discussed in Sec. IV the t sums converge for any finite volume provided that the quantity

$$\sum_{\vec{p}\neq\vec{p}'} f_{\vec{p}\vec{p}'} = \sum_{\vec{p}\neq\vec{p}'} \sum_{ij} \left[\exp\left(-\frac{\pi^2}{g^2} t^i_{\vec{p}} V^{ij}_{\vec{p}\vec{p}'} t^j_{\vec{p}'}\right) - 1\right] \tag{5.17}$$

does not grow more rapidly than V for large V. This requirement of clustering is satisfied here (unlike in the two-dimensional case discussed in Sec. IV), since from (5.16) we see that, for large separations $r = |\vec{p} - \vec{p}'|$ and $\gamma_{\vec{k}}$ given by (5.10),

$$V^{ij}_{\vec{p}\vec{p}'} \propto \frac{\cos\pi r}{r^2}\left[\delta_{ij} - \left(\frac{r_i r_j}{r^2}\right)\right]. \tag{5.18}$$

Having argued that the restricted t sum, and hence the N sum, is convergent we can now examine individual terms to see whether there is any way to generate a contribution to (5.8) which grows linearly with the separation of the charges. The convergence of the N sum means it is dominated by configurations with widely separated loops of nonvanishing $\vec{N}_{\vec{p}}$. Their contribution to the change in energy between the configuration with no charges and that with charges separated by a distance D is dominated by the term proportional to

$$\delta E = \exp\left[-\left(\frac{K}{g^2}\right)\sum_{\substack{\text{loops}\\ i}} t_i^2 P_i\right] \prod_{\substack{\text{loops}\\ i}} \left[\cos\left(2\pi \sum_{\substack{\text{loop}\\ \text{surface}}} \vec{t}_i \cdot \vec{E}^{\text{Coul}}\right) - 1\right], \tag{5.19}$$

where K is some (large) constant and P_i is the perimeter of the ith loop. Now consider what happens to this quantity as the distance D between the charges is increased. For any fixed loop the contribution decreases as $1/D^2$ as the charges recede. Therefore the contribution to the energy from such configurations also decreases with increasing D at least as fast as $1/D^2$. In addition one obtains contributions to the energy which are independent of D from loops of fixed size which are close to either individual charge and remain so as D changes; these are of course nothing but self-energy contributions and their magnitude is suppressed by a factor $e^{-K/g^2} \ll 1$. The only way to obtain an increase in the argument of any of the cosines in (5.19) which is proportional to D is to let the area of that loop increase at least linearly in D as D is increased. However, this can only be achieved by increasing the perimeter[22] at least as fast as $D^{1/2}$. Hence the contribution to the energy from such a term is exponentially damped with increasing D. It follows that one cannot identify any class of N-loops whose contribution grows linearly with D. Thus we conclude that $\delta E(\{\epsilon\})$ at most gives a negligible correction $O(e^{-K/g^2})$ to the Coulomb energy.

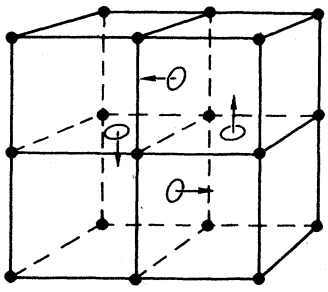

FIG. 10. The simplest nonvanishing $\vec{N}_{\vec{p}}$ distribution which satisfies $(\vec{\nabla}\cdot\vec{N})_{\vec{p}} = 0$.

We stress once again that this argument, which is made by examining the N sum term by term, is invalid in the strong-coupling limit, where the summation is not expected to converge. We have not computed the value of g^2 beyond which this series diverges, nor have we studied the nature of the phase transition from unconfined QED for $g^2 \ll 1$ to the confined phase when $g^2 \gg 1$.

A more physical picture of the difference between two and three spatial dimensions can be ob-

tained using the strong-coupling limit as a starting point. We found there that the ground state of the system consisted of a static coherent cloud of transverse photons, described by the variables $\vec{L}_{\vec{p}}$. This cloud serves to cancel the Coulomb field in all of space except along the line between the charges; we may describe this phenomenon as a focusing of the flux. As the coupling decreases quantum fluctuations in the variables $\vec{L}_{\vec{p}}$ become more important. The periodicity of the potential means that these fluctuations change the eigenvalues of the $\vec{L}_{\vec{p}}$ by integers. Our calculation is a study of whether or not the fluctuations can randomize the configuration sufficiently to completely destroy the coherence. In a two-dimensional lattice the vectors $\vec{L}_{\vec{p}}$ are constrained to lie perpendicular to the plane of the lattice. They each can fluctuate in magnitude by integer amounts, but fluctuations cannot destroy the fact that they are aligned, nor give a vanishing $\langle \vec{L}_{\vec{p}} \rangle$ for any $0 < |\epsilon_{\vec{p}}| \leq \frac{1}{2}$. However, in three dimensions fluctuations can cause the vector $\vec{L}_{\vec{p}}$ to rotate as the individual components change in magnitude by integers. This new degree of freedom allows the coherence of the state to be completely eliminated, even though the $\langle |\vec{L}_{\vec{p}}| \rangle$ are still nonvanishing.

VI. SUMMARY AND CONCLUSIONS

We have developed and applied a Hamiltonian variational approach to the study of QED, formulated both compactly and noncompactly, on a (spatial) lattice in both 2+1 and 3+1 dimensions. In the course of obtaining results previously reported by others,[3,4] we have introduced a different language and physical picture of confinement for these models.

The important features of our method are

(i) explicit separation of classical and quantum variables,

(ii) imposition of the appropriate periodic boundary conditions on the variational wave function in the compact formulation for which the potential is periodic, and

(iii) factorization of the compact problem, in the weak-coupling limit, into a variational calculation which is the same as that of the noncompact case, plus a statistical-mechanics-like calculation of the corrections from nonperturbative effects—these are the N sums.

In the compact formulation of the theory step (i) introduces not only the static external charges and resultant classical Coulomb field, but an additional set of classical variables: the $\epsilon^a_{\vec{p}}$. The existence of these is a consequence of the periodicity of the Hamiltonian, which means it is invariant under translations of B by multiples of 2π.

In the absence of charges the $\epsilon^a_{\vec{p}}$ are taken to be zero corresponding to zero mean electric field strength. Charges are introduced using the gauge-invariant operator $\bar{\psi} e^{-iA} \psi$, which creates charges joined by a line of unit flux. Reinterpreting this state as a static Coulomb field plus the quantum excitation of a coherent photon state gives nonvanishing $\epsilon^a_{\vec{p}}$ over all of space. In the sector of the Hilbert space defined by these $\epsilon^a_{\vec{p}}$, the total flux on any link is given by the sum of the quantum excitations plus the Coulomb field and is hence an integer for every eigenstate. In the strong-coupling limit the coherent state is an eigenstate of the system, and in the ground state the flux is localized along the shortest path between the two charges. As the coupling decreases fluctuations become more and more important; they depend on the $\epsilon^a_{\vec{p}}$ and hence are different in the presence of charges, which polarize the vacuum, than they are in the charge-free sector. The dependence of the fluctuations upon the $\epsilon^a_{\vec{p}}$ arises entirely in the nonperturbative effects which correspond to tunnelings between different minima of the potential—these effects are included in our approach by the introduction of the properly periodic wave functions in step (ii). In Euclidean path integral calculations these effects are included in the nontrivial classical (or semiclassical) solutions—the N's of our N sums give rise to the same effects as the monopoles of Polyakov in 2+1 dimensions that lead to linear confinement in the weak- as well as strong-coupling limit. In 3+1 dimensions they give rise to the same effects as the loops of monopoles and do not lead to confinement in the weak-coupling limit.

Much of the physics we have observed in this calculation was anticipated and qualitatively described by Polyakov,[3] with further support from the work of Banks *et al.*[4] Our method of calculation, although quite different, is found to be capable of reproducing the qualitative effects which arise in path-integral calculations from nontrivial classical or semiclassical solutions. It also highlights the crucial importance of the photon self-interactions that are introduced by the compact formulation. Further, we make quite different approximations from those made by previous authors and yet arrive at the same conclusions, thus strengthening our belief that these conclusions are correct, and not just the consequences of some simplifying assumption made during the calculation. We make no attempt here to obtain precise energies, or to investigate the intermediate coupling range—for example to find the value of the critical coupling in the (3+1)-dimensional theory. We are interested at this stage only in certain gross qualitative statements about the theory. However, the demonstrated power of the recursive

variational method invites further study of such questions by these methods.

Much still remains to be learned about lattice QED even on a qualitative level. The introduction of quantum fermions is an obvious and interesting next step. We are now optimistic that these methods can also be extended to examine the confining properties of non-Abelian gauge theories on a lattice, in all regions of coupling strength.

ACKNOWLEDGMENTS

We thank our colleagues at SLAC, especially Stephen Shenker, for many helpful discussions. One of us (H. R. Q.) wishes to thank the Alfred P. Sloan foundation for financial support and the Physics Department at Stanford University for hospitality during the course of this work. This work was supported by the Department of Energy.

[1]S. D. Drell, M. Weinstein, and S. Yankielowicz, Phys. Rev. D 14, 487 (1976); 16, 1769 (1977).

[2]K. G. Wilson, Phys. Rev. D 10, 2445 (1974); J. Kogut and L. Susskind, *ibid.* 11, 395 (1975); L. Susskind, in *Trends in Elementary Particle Theory*, proceedings of the International Summer Institute for Theoretical Physics, Bonn, 1974, edited by H. Rollnik and K. Dietz (Springer, New York, 1975), p. 234; T. Banks, J. Kogut, and L. Susskind, Phys. Rev. D 13, 1043 (1976); K. G. Wilson, in Renormalization Theory, proceedings of the International School of Mathematical Physics, "Ettore Majorana," Erice, Italy, 1975, edited by G. Velo and A. Wightman (D. Reidel, Dordrecht, Holland, Boston, 1976). Related discussions of these kinds of models are also given in B. E. Baaquie, Phys. Rev. D 16, 3040 (1977) and in J. F. Willemsen, ibid. 17, 2082 (1978).

[3]A. M. Polyakov, Phys. Lett. 59B, 82 (1975); Nucl. Phys. B120, 429 (1977).

[4]T. Banks, R. Myerson, and J. Kogut, Nucl. Phys. B129, 493 (1977); J. Villain, J. Phys. 36, 581 (1975). See also M. Peskin, Ann. Phys. (N.Y.) 113, 122 (1978).

[5]The potential $(1/e^2a^2)\sum_{\vec{p}}(1-\cos ea^2B_{\vec{p}})$ arises from the product around each plaquette of four link operators of the type $U_{\vec{p}}^{\hat{b}} \equiv e^{-iea^2A_{\vec{p}}^{\hat{b}}}$. This product has the property that it is gauge invariant. In a theory with charges, combinations of the form $\psi_{\vec{p}}^{\dagger}U_{\vec{p}}^{\hat{b}}\psi_{\vec{p}+\hat{b}}$ are also gauge invariant. U is the lattice equivalent of the Schwinger line integral.

[6]The lattice operator ∇^2 is defined by (2.19) for points $\vec{p}$ in the interior of the lattice; if $\vec{p}$ lies on the edge the definition is modified. A general definition is given by $(\nabla^2\phi)_{\vec{p}} = \sum_{\text{nearest neighbors}} \phi_{\vec{p}'} - (\text{number of nearest neighbors})\times\phi_{\vec{p}}$. This is equivalent to requiring Neumann boundary conditions for $\vec{E}$ in the continuum theory.

[7]This follows from (2.9)–(2.11) with $\Lambda_{\vec{p}}$ set equal to the constant function.

[8]In two dimensions for any finite lattice the values $L_{\vec{p}}$ corresponding to any given $\vec{E}_{\vec{p}}$ are uniquely defined, if we choose to define $L_{\vec{p}} \equiv 0$ for $\vec{p}$ outside the lattice. An alternate formulation with periodic boundary conditions leaves the $L_{\vec{p}}$ undefined by an arbitrary ($\vec{p}$-independent) constant. In three dimensions we will see that $\vec{L}_{\vec{p}}$ is defined by (2.20) only up to $\nabla\sigma_{\vec{p}}$ where $\sigma_{\vec{p}}$ is an arbitrary integer for each point $\vec{p}$.

[9]The appearance of fractional ϵ variables here is a generalization of the property found in the (1+1)-dimensional Higgs model [H. R. Quinn and M. Weinstein, Phys. Rev. D 17, 1063 (1978)]. There we found a single global ϵ parameter which could be identified with the θ parameter of the instanton picture.

[10]F. Bloch, Z. Phys. 52, 555 (1928). See also R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, Mass., 1965), Vol. III, Chap. 13.

[11]S. D. Drell, M. Weinstein, and S. Yankielowicz, Phys. Rev. D 14, 1627 (1976).

[12]It is obvious that this ansatz will grossly underestimate the coefficient of $\cos(2\pi\epsilon)$. Our result depends on the overlap of the wave function tails under the barrier, which is just where the Gaussian approximation is poorest. One could improve the estimate by choosing a trial form which is exponentially damped in this region, rather than Gaussian, for example $\psi_{tr}(\theta) = \exp(-\alpha_0\theta\tanh\beta_0\theta)$. We use the Gaussian form in order to be able to make simple analytic estimates, since we are here interested only in the qualitative features of the result rather than in making detailed quantitative statements. Since our emphasis is to show that the ϵ-dependent effects do not vanish, even a gross underestimate is sufficient. We remark also that our Gaussian wave function would be the best trial form if we were to study the Villain approximation to the Hamiltonian, which replaces the periodic potential $(1/g^2)(1-\cos B)$ by $V(B) = (1/2g^2)(B-2\pi n)^2$ for each segment $2\pi(n-\frac{1}{2}) \leq B \leq 2\pi(n+\frac{1}{2})$. Hence, it is not surprising that our treatment yields similar results to those obtained using this Villain potential in a path-integral calculation (see Ref. 4).

[13]The factor $(2\pi\gamma)^{-1/2}$ in the normalization of $\chi(\theta)$ tells us that in the $g\to\infty$ limit the N sums do not converge, though ratios such as $\langle\chi|H|\chi\rangle/\langle\chi|\chi\rangle$ do.

[14]With periodic boundary conditions ($E^y_{\vec{p}_x,N_0} = E^y_{\vec{p}_x,-N_0-1}$, $E^x_{N_0,\vec{p}_y} = E^x_{-N_0-1,\vec{p}_y}$) the variable $\sum_{\vec{p}}L_{\vec{p}}$ does not enter the Hamiltonian and the ground state is a superposition of all states $|m_{\vec{p}} = \text{constant}\rangle$,

$$\psi(\{\theta\}) = \sum_m \exp(im\sum_{\vec{p}}\theta_{\vec{p}}).$$

[15]The absence of any singular behavior in the theory as $g^2\to\infty$ is indicated by the existence of an apparently convergent perturbation expansion about this limit [see for instance Banks *et al.*, Ref. 2].

[16]See, for example, K. Huang, *Statistical Mechanics* (Wiley, New York, 1963); R. P. Feynman, *Statistical Mechanics* (Benjamin, Reading, Mass., 1972).

[17]We thank Stephan Shenker for reminding us of this identity and suggesting its utility in this problem.

[18]It is clear from (4.23) that, given any distribution of nonvanishing $N_{\vec{p}}$, the maximum contribution to Z for that distribution comes from all these $N_{\vec{p}} = \pm 1$. Terms with larger values of $|N_{\vec{p}}|$ are suppressed relative to these by powers of e^{-1/g^2}. We leave further demonstration of the validity of the truncation as an exercise for the reader.

[19]The form of (4.40) is effectively a "time slice" of the problem solved by Polyakov in evaluating the Wilson loop integral. The approximations we make from this point on destroy the explicit periodicity in $\epsilon_{\vec{p}}$; they are reasonable only if we remember that each $\epsilon_{\vec{p}}$ is defined to lie in the range $-\frac{1}{2} \leq \epsilon_{\vec{p}} \leq \frac{1}{2}$.

[20]Once again the ground state is $(\nabla \times m) = 0$, but $\vec{\nabla} \cdot \vec{m}$ is arbitrary. See Ref. 14.

[21]The restriction on the t sum could readily be removed, since so doing is equivalent to multiplication of numerator and denominator in (5.14) by a (divergent) constant. However, we have not found this to be a useful way to proceed.

[22]The restriction on the t sum in (5.14) excludes changes in the surface which do not alter the perimeter. By Gauss's law all surfaces which have the same perimeter have the same $\sum_{\text{surface}} \vec{t} \cdot \vec{E}$ modulo integers which would not alter (5.19).